

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: August 27, 2004, 16:27:32 ; Search time 474 Seconds
(without alignments)
2948.644 Million cell updates/sec

Title: US-10-081-775-2

Perfect score: 1718

Sequence: 1 MSTLGHNMESPHHTDPS.....RKRVVRVFSQGMGKASE 329

Scoring table:

BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 3373863 seqs, 2124099041 residues

Total number of hits satisfying chosen parameters: 6747726

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+ p2n.model -DEV=xlh
-Q=/cgn2.1/USPTO.spool/US10081775/runat 20082004.170212.11117/app_query.fasta_1.519
-DB=N Geneseq 29Jan04 -QFMT=fastap -SUFFIX=ring -MINMATCH=0.1 -LOOPCL=0
-LOOPEXT=0 -UNITS=bits -START=1 -END=-1 -MATRIX=BLOSUM62 -TRANS=human40.cdi
-LIST=45 -DOALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15
-MODE=LOCAL -OUTFMT=pct -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US10081775.0CGN 1 1 470 @runat 20082004.170212.11117 -NCPU=6 -ICPU=3
-NO MMAP -LARGQUERY -NEG SCORES=0 -WAIT -DSBLOCK=100 -LONGLOG
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : N Geneseq 29Jan04.*

- 1: Geneseq1980s.*
- 2: Geneseq1990s.*
- 3: Geneseq2000s.*
- 4: Geneseq2001as.*
- 5: Geneseq2001bs.*
- 6: Geneseq2002s.*
- 7: Geneseq2003as.*
- 8: Geneseq2003bs.*
- 9: Geneseq2003cs.*
- 10: Geneseq2004s.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1718	100.0	1567	ACD91434	Acd91434 Human G-p
2	1691	98.4	990	AAS42339	Aas42339 Human cDN
3	1691	98.4	990	ABZ43046	Abz43046 Human GPC
4	1691	98.4	998	ABK37625	Abk37625 DNA encod
5	1691	98.4	998	ABQ88371	Abq88371 Human G p
6	1691	98.4	998	ABQ88372	Abq88372 Human G p
7	1691	98.4	1390	ADC86344	Adc86344 Human GPC
8	1649	96.0	963	AAH31850	Aah31850 Human olf

9	1649	96.0	966	6	ABK16633	Abk16633 Human G-c
10	1649	96.0	966	6	ABK68612	Abk68612 Human DNA
11	1376	80.1	948	7	ABZ77974	Abz77974 Human G p
12	978.5	57.0	1300	9	ADC86380	Adc86380 Human GPC
13	950.5	55.3	948	7	ABZ77971	Abz77971 Human G p
14	931.5	54.2	942	4	AAH31721	Aah31721 Human olf
15	931.5	54.2	985	4	AAH32052	Aah32052 Human olf
16	931.5	54.2	990	6	ABZ43027	Abz43027 Human GPC
17	931.5	54.2	990	6	ABK68587	Abk68587 Human GPC
18	931.5	54.2	1012	6	ABSS8762	Abss8762 Human DNA
19	931.5	54.2	1114	6	AAF88624	Aaf88624 Human G-p
20	931.5	54.2	1390	9	ADC86366	Adc86366 Human GCR
21	920	53.6	1076	4	AAD12945	Aad12945 Human G-p
22	909.5	52.9	1012	6	ABSS9230	Abss9230 Human G-p
23	907	52.8	945	4	ABA81539	Abas81539 Human GPC
24	904	52.6	933	4	AAH31738	Aah31738 Human olf
25	904	52.6	936	5	AAS42265	Aas42265 Human cDN
26	904	52.6	936	6	ABZ42976	Abz42976 Human GPC
27	904	52.6	936	6	ABK68617	Abk68617 Human DNA
28	904	52.6	936	6	ABK37551	Abk37551 DNA encod
29	904	52.6	985	4	ABA81537	Abas81537 Human GPC
30	904	52.6	1001	5	AAS15906	Aas15906 DNA encod
31	904	52.6	1336	9	ADC86330	Adc86330 Human GPC
32	903	52.6	947	4	ABA81538	Abas81538 Human GPC
33	901.5	52.5	1013	6	ABQ88366	Abq88366 Human G p
34	899.5	52.4	1360	9	ADC86368	Adc86368 Human GPC
35	889.5	51.8	954	4	AAH32330	Aah32330 Human olf
36	889.5	51.8	957	5	AAS42259	Aas42259 Human cDN
37	889.5	51.8	957	6	ABZ42968	Abz42968 Human GPC
38	889.5	51.8	957	6	ABK68605	Abk68605 Human DNA
39	889.5	51.8	957	6	ABK37545	Abk37545 DNA encod
40	888.5	51.7	980	6	ABQ88354	Abq88354 Human G p
41	888.5	51.7	980	6	ABQ88355	Abq88355 Human G p
42	886.5	51.6	982	4	AAD19141	Aad19141 Human G-p
43	886.5	51.6	1592	6	ABSS1301	Abss1301 Human REM
44	885.5	51.5	963	4	AAH31827	Aah31827 Human olf
45	885.5	51.5	966	5	AAS42338	Aas42338 Human cDN

ALIGNMENTS

RESULT 1

ACD91434

ID ACD91434 standard; cDNA; 1567 BP.

XX AC ACD91434;

XX AC ACD91434;

DT 22-SEP-2003 (first entry)

XX Human G-protein coupled receptor HGPRBMV25 cDNA.

Human; G-protein coupled receptor; HGPRBMV25; immune disease; inflammatory disease; arthritis; asthma; AIDS; psoriasis; graft-versus-host disease; systemic lupus erythematosus; reproductive disorder; varicocele; orchitis; neural disorder; Alzheimer's disease; Parkinson's disease; depression; schizophrenia; cardiovascular disorder; hypertension; acute heart failure; pulmonary disorder; endocrine disorder; obesity; diabetes; anorexia; bone disorder; osteoporosis; pain; cancer; chromosome identification; gene therapy; gene; ss.

XX Homo sapiens.

OS US2003060409-A1.

XX 27-MAR-2003.

XX 21-FEB-2002; 2002US-00081775.

XX 21-FEB-2001; 2001US-0270134P.

XX 27-MAR-2001; 2001US-0278952P.

XX (RAMA/) RAMANATHAN C S.

PA

957	Db	CGCTATGCCACAACTCTCACTGACACCATCAATTGCCACATAGGGGTGGCAGCTGTAGTG	101
161	Qy	ArgGlySerLeuLeuMetLeuProCysProPhePheIleGlyArgLeuAsnPheCysGln	180
1017	Db	CGAGGCTCCCTGCTCATGTCTCCATGTCCCTCTTTATTTGGGGTTTGAACTTCTGCGCAA	1076
181	Qy	SerHisValIleLeuHisThrTVrCysGluHisMetAlaValValllysLeuAlaCysGly	200
1077	Db	AGCCATGTATCTCATCACAGCTACTGTGTAGCACAATGGCTGTGGTGAAGCTGGCCTGTGGA	1136
201	Qy	AspThrArgProAsnArgValTVrGlyLeuThrAlaAlaLeuLeuValIleGlyValAsp	220
1137	Db	GACACAGGCCTAACCGTGTGTATGGCTGACAGCTGCACTGTGTGGTCATTTGGGGTTGAC	1196
221	Qy	LeuPheCysIleGlyLeuSerTVrAlaLeuIleAlaGlnAlaValLeuArgLeuSerSer	240
1197	Db	TTGTTTTTGCAATTTGGTCTCTCCTATGGCCTAATTGCACAAGCTGTCTTCGCTCTCATCC	1255
241	Qy	HisGluAlaArgSerLysAlaLeuGlyThrCysGlySerHisValCysValIleLeuIle	260
1257	Db	CATGAAGCTCGTGCCAAAGGCCCTAGGACCTGTGTGTTCCCATGTCCTGTGTATCCTCATC	1311
261	Qy	SerTyrThrProAlaLeuPheSerPheThrHisArgPheGlyHisHisValProVal	280
1317	Db	TCCTATACACAGCCCTCTCTCCTTTTTTACACACCCGCTTTGGCCATCACGTTCCAGTC	1371
281	Qy	HisIleHisIleLeuLeuAlaAsnValTVrLeuLeuLeuProProAlaLeuAsnProVal	300
1377	Db	CATATTCACATTTCTTTTGGCCAAATGTTTATCTGCTTTTGGCCACTGCTTTAATCCTGTG	1433
301	Qy	ValTVrGlyValLysThrLysGlnIleArgLysArgValValArgValPheGlnSerGly	320
1437	Db	GTAATGAGTTTAAAGCAACACAGATCCGTAAACAGATTGTCCAGGTTGTTTCAAAAGTGGG	1497
321	Qy	GlnGlyMetGlyIleLysAlaSerGlu	329
1497	Db	CAGGGAATGGGCATCAAGGCATCTGAG	1523
RESULT 2			
AAS42339			
ID	AAS42339 standard; cDNA; 990 BP.		
XX	AAS42339;		
AC	AAS42339;		
DT	18-DEC-2001 (first entry)		
XX	Human cDNA encoding olfactory receptor AOLFRI41.		
XX	Human; olfactory receptor; G protein-coupled receptor; GPCR; odourant;		
KW	ss; food additive; cosmetic; fragrance; pharmaceutical additive.		
XX	Homo sapiens.		
OS	WO200168805-A2.		
XX	20-SEP-2001.		
PN	13-MAR-2001; 2001WO-US007771.		
XX	13-MAR-2000; 2000US-0188914P.		
PR	24-MAR-2000; 2000US-0192033P.		
PR	12-APR-2000; 2000US-0198474P.		
PR	24-APR-2000; 2000US-0199335P.		
PR	26-MAY-2000; 2000US-0207702P.		
PR	23-JUN-2000; 2000US-0213849P.		
PR	16-AUG-2000; 2000US-0226534P.		
PR	07-SEP-2000; 2000US-0230732P.		
PR	07-FEB-2001; 2001US-0266862P.		
XX	(SENO-) SENOMYX INC.		
PA	Zozulya S;		
PI			

(FEDE/) FEDER J N.
(MINT/) MINTIER G A.

Ramanathan CS, Feder JN, Mintier GA;
WPI: 2003-521919/49.
P-PSDB; AB042809.

New nucleic acid molecule encoding a human G-protein coupled receptor (HGPRBMV25) is useful for diagnosing, preventing or treating diseases involving the receptor, e.g. inflammation, diabetes, asthma, hypertension or cancer.

Claim 1; Fig 1A-B; 139pp; English.

The invention describes an isolated nucleic acid molecule comprising a sequence that is at least 95% identical to a polynucleotide encoding novel human G-protein coupled receptor HGPRBMV25. The nucleic acid molecule, polypeptide and antibody are useful in diagnosing, preventing, treating or ameliorating medical conditions where GPCR is directly or indirectly involved, such as immune or inflammatory diseases (e.g. arthritis, asthma, AIDS, graft-versus-host disease, psoriasis or systemic lupus erythematosus), reproductive disorders (e.g. varicocele or orchitis), neural disorders (e.g. Alzheimer's disease, Parkinson's disease, depression or schizophrenia), cardiovascular disorders (e.g. hypertension or acute heart failure), pulmonary disorders, endocrine disorders (e.g. obesity, diabetes or anorexia), bone disorders (e.g. osteoporosis), pain or cancer. The polynucleotide may also be used in chromosome identification, in identifying organisms from minute biological samples, or as molecular weight markers. This sequence encodes a novel human G-protein coupled receptor HGPRBMV25

Sequence 1567 BP; 367 A; 382 C; 314 G; 504 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.: 1,32e-171 Length: 1567
Score: 1718.00 Matches: 329
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 8 Gaps: 0

US-10-081-775-2 (1-329) x ACD91434 (1-1567)

QY 1 MetSerSerThrLeuGlyHisAsnMetGluserProHisHisThrAspValaspProSer 20
Db 537 AVGTCCAGACACTCTTGGGCCAACCAATGNAATCTCTCATCACACTGATGTTGACCCTTCT 596
QY 21 ValPhePheLeuGlyIleProGlyLeuGluGlnPheHisLeuTrpleuSerleuPro 40
Db 597 GTCTTCCTCCTCGGCATCCAGCTCTGGAAACAATTTCATTTGTGGCTCTCATCTCCCT 656
QY 41 ValCysGlyLeuGlyThrAlaThrIleValGlyAsnIleThrIleLeuValValAla 60
Db 657 GTGTGTGGCTTAGGCACACCAAAVTGGGGAATAATDACTATTCTGTGTTGTTGGC 716
QY 61 ThrGluProValLeuHisLysProValTyrLeuPheLeuCysMetleuSerThrIleAsp 80
Db 717 ACTGAACCAAGTCTTGCAAGCTGTGTACCTTTTCTGTGCATGCTCTCAACCATCGAC 776
QY 81 LeuAlaAlaSerValserThrValProLysLeuLeuAlaIlePheTrpCysGlyAlaGly 100
Db 777 TTGGCTGCCTCTCTCTCCACAGTTCACAGCTACTGGCTATCTCTGGGTGTGGAGCCGA 836
QY 101 HisIleSerAlaSerAlaCysLeuAlaHisMetPheIleHisAlaPheCysMetMet 120
Db 837 CATATATCTCGCCTTCGCTGGCAGCATATGTTCTTATTTCATTCGCTTCTCATGATG 896
QY 121 GluSerThrValLeuLeuAlaMetAlaPheAspArgTyrValAlaIleCysHisProLeu 140
Db 897 GAGTCCACTGTGCTACTGGCCATGGCCCTTTGATCGCTACGTGGCCATCTGCCACCTC 956
Db 956

XX WPI: 2001-570867/64.
 DR P-PSDB; AAU24646.
 XX

PT Nucleic acids encoding human olfactory G protein-coupled receptors,
 PT useful for screening for compounds involved in olfactory sensation, where
 PT the compounds can be used in the food, pharmaceutical and cosmetic
 PT industries to customize odors.
 XX

PS Claim 1; Page 142; 319pp; English.
 XX

CC The invention relates to nucleic acids encoding human olfactory
 CC receptors, OR, (a G protein-coupled receptor, GPCR). The OR's
 CC specifically recognise molecules, odourants, that elicit specific
 CC olfactory sensation. The human olfactory receptors and polynucleotides
 CC encoding them are useful for screening a library of chemical compounds
 CC for compounds that are involved in olfactory sensation. Modulators of
 CC their activity are useful for pharmacological and genetic modulation of
 CC olfactory signalling pathways. Therefore, they can be used in the food,
 CC pharmaceutical and cosmetic industries to customise odours and
 CC fragrances. The present sequence encodes a human olfactory receptor of
 CC the invention
 XX

SQ Sequence 990 BP; 188 A; 285 C; 221 G; 296 T; 0 U; 0 Other;
 XX

Alignment Scores:
 Pred. No.: 4.97e-169 Length: 990
 Score: 1691.00 Matches: 325
 Percent Similarity: 99.09% Conservative: 1
 Best Local Similarity: 98.78% Mismatches: 3
 Query Match: 98.43% Indels: 0
 DB: 5 Gaps: 0

US-10-081-775-2 (1-329) x AAS42339 (1-990)

QY 1 MetSerThrLeuGlyHisAsnMetGluSerProHisHisThrAspValAspProSer 20
 DB 1 ATGTCCAGCACTCTTGGCCCAACATGGAATCTCTTAATACACACTGATGTTGACCCCTCT 60
 QY 21 ValPheLeuLeuGlyVileProGlyLeuGluGlnPheHisLeuTrpLeuSerLeuPro 40
 DB 61 GTCTTCTTCCCTGGGATCCAGGCTCGAACAATTTCAATTTGGCTCTCACTCCCT 120
 QY 41 ValCysGlyLeuGlyThrAlaThrIleValGlyAsnIleThrIleLeuValValAla 60
 DB 121 GTGTGTGGCTTAGGCACAGGCACAAATGTGGCAATATACTATCTGTTGTTGTTGCC 180
 QY 61 ThrGluProValLeuHisLysProValTyrLeuPheLeuCysMetLeuSerThrIleAsp 80
 DB 181 ACTGAACAGTCTTGCAAGCCCTGTGTACCTTTTCTGTGCATGCTCTCAACCATCGAC 240
 QY 81 LeuAlaAlaSerValSerThrValProLysLeuLeuAlaIlePheTrpCysGlyAlaGly 100
 DB 241 TTGGCTGCCTCTGTCTCCACAGTTCACAGTCTCCAGCTACTGGCTATCTTCTGGTGAGCCGA 300
 QY 101 HisIleSerAlaSerAlaCysLeuAlaHisMetPheIleHisAlaPheCysMetMet 120
 DB 301 CATATATCTGCCCTTGCTGCTGGCCACAGATGTTCTTCATTCATGCTTCTGCATGATG 360
 QY 121 GluSerThrValLeuLeuAlaMetAlaPheAspArgTyrValAlaIleCysHisProLeu 140
 DB 361 GAGTCCACTGTGTACTGGCCATGGCTTTGTAGCTGCTACGTGGCCATCTGCCACCCACTC 420
 QY 141 ArgTyrAlaThrIleLeuThrAspThrIleIleAlaHisIleGlyValAlaValVal 160
 DB 421 CGTATGCCACATCTCTACTGACACCATCATTTGCCCATAGGGGTGGCAGTGTAGTG 480
 QY 161 ArgGlySerLeuLeuMetLeuProCysProPheIleGlyArgLeuAsnPheCysGln 180
 DB 481 CGAGGCTCCCTGCTCATGCTCCATGCTCCCTCTTCTTATTTGGGGCTTTGAACTTCTGCCAA 540
 QY 181 SerHisValIleLeuHisThrTyrCysGluHisMetAlaValValLysLeuAlaCysGly 200

541 AGCCATGTGATCCTACACAGCTACTGTGAGCACATGCTGTGGTGAAGCTGGCCCTGTGGA 600
 201 AspThrArgProAsnArgValTyrGlyLeuThrAlaAlaLeuLeuValIleGlyValAsp 220
 601 GACACAGGCTAACCGTGTGTATGGCTGACAGTCACTGTGTGCTCATTTGGGTTGAC 660
 221 LeuPheCysIleGlyLeuSerTyrAlaLeuIleAlaGlnAlaValLeuArgLeuSerSer 240
 661 TTGTTTGGCATGTGCTCTCTCTATGCCCTAAGTGCAACAGCTGCTCTTCCCTCTCATCC 720
 241 HisGluAlaArgSerLysAlaLeuGlyThrCysGlySerHisValCysValIleLeuIle 260
 721 CATGAAGCTCGGTCCAGGCCCTTAGGAGCTGTGGTTCCTCATGCTGTGTGTCATCTCATC 780
 261 SerTyrThrProAlaLeuPheSerPhePheThrHisArgPheGlyHisHisValProVal 280
 781 TCTTATACACAGCCCTCTCTCTCTTTTACACACCGCTTTGGCCATCACGTTCCAGTC 840
 281 HisIleHisIleLeuLeuAlaAsnValTyrLeuLeuProProAlaLeuAsnProVal 300
 841 CATATTCACTCTTTTGGCCAATGTTTATCTGCTTTTCCACCTGCTCTTAACTCTCTGTG 900
 301 ValTyrGlyValLysThrLysGlnIleArgLysArgValValArgValPheGlnSerGly 320
 901 GTATATGGAGTTAAGCAACAGATCCGTAAAGAGTTGTCCAGGTTGTTTCAGAGTGGG 960
 321 GlnGlyMetGlyVileLysAlaSerGlu 329
 961 CAGGGAATGGGCATCAAGGCATCTGAG 987

RESULT 3
 ABZ43046
 ID ABZ43046 standard; DNA; 990 BP.
 XX
 AC ABZ43046;
 XX
 DT 06-MAR-2003 (first entry)
 XX
 DE Human GPCR polynucleotide SEQ ID NO 353.
 XX
 KW Human; GPCR; G protein coupled receptor; signal transduction; olfactory;
 KW drug development; gustatory; taste; fragrance; gene; ds.
 XX
 OS Homo sapiens.
 XX
 FN W0200216548-A2.
 XX
 PD 28-FEB-2002.
 XX
 PF 30-JUL-2001; 2001WO-IB001446.
 XX
 PR 04-AUG-2000; 2000JP-00237818.
 PR 13-FEB-2001; 2001JP-00034434.
 XX
 PA (NISC-) JAPAN SCI & TECHNOLOGY CORP.
 XX
 PI Haga T, Takeda S, Mitaku S;
 XX
 WPI: 2002-304118/34.
 DR P-PSDB; ABP95772.
 XX
 PT Database global search for G protein-coupled receptors, proteins and
 PT identifying targets for drug development.
 XX
 PS Claim 9; SEQ ID NO 353; 97pp + Sequence Listing; Japanese.
 XX
 CC The invention relates to a method for screening G protein-coupled
 CC receptor (GPCR) genes (ABZ42870-ABZ43216) and/or GPCR proteins (ABP95596-
 CC ABP95942) by extracting open-reading frames containing 6-8 transmembrane
 CC domains with 250-1000 amino acid residues to give a gene homologous with
 CC a known GPCR gene. The receptor proteins and encoded genes are useful for
 CC studying in vivo signal transduction mechanism and identifying targets

CC for drug development e.g. based on olfactory and gustatory receptors in
 CC form of agonists and antagonists by screening intrinsic and extrinsic
 CC ligands as bitter taste inhibitors, taste enhancers and fragrance
 CC improvers. Note: The sequence data for this patent did not form part of
 CC the printed specification, but was obtained in electronic format directly
 CC from WIPO at ftp.wipo.int/pub/published_pct_sequences
 XX
 SQ Sequence 990 BP; 188 A; 285 C; 221 G; 296 T; 0 U; 0 Other;

Alignment Scores:
 Pred. No.: 4,97e-169 Length: 990
 Score: 1691.00 Matches: 325
 Percent Similarity: 99.09% Conservative: 1
 Best Local Similarity: 98.78% Indels: 0
 Query Match: 98.43% Gaps: 0
 DB:

US-10-081-775-2 (1-329) x ABZ43046 (1-990)

QY 1 MetSerThrLeuGlyHisAsnMetGluSerProHisHisThrAspValAspProSer 20
 DB 1 ATGTCAGACACTCTGGCCACACATGGAATCTCTTAATCACATGATGTTGACCCCTCT 60
 QY 21 ValPhePheLeuLeuGlyIleProGlyLeuGluGlnPheHisLeuTrpLeuSerLeuPro 40
 DB 61 GTCTTCTCTCTGGGCATCCAGGCTCGAACAATTTCAATTTGTGGCTCTCACTCCCT 120
 QY 41 ValCysGlyLeuGlyThrAlaThrIleValGlyAsnIleThrIleLeuValValAla 60
 DB 121 GTGTGGCTGTAGGCACAGCCACAAATTTGGGCAATATAAATTAATTTCTGTTGTTGCC 180
 QY 61 ThrGluProValLeuHisIleProValTyrLeuPheLeuCysMetLeuSerThrIleAsp 80
 DB 181 ACTGAACAGACTTGCACAAAGCCTGTGATCTTTCTGTGCATGCTCTCAACCATCGAC 240
 QY 81 LeuAlaAlaSerValSerThrValProIysLeuLeuAlaIlePheTrpCysGlyAlaGly 100
 DB 241 TTGGCTGCTCTGTCTCCACAGTCTCCAGCTTCCAGCTACTGGCTATCTTGTGTGGAGCCGA 300
 QY 101 HisIleSerAlaSerAlaCysLeuAlaHisMetPhePheIleHisAlaPheCysMetMet 120
 DB 301 CATATATCTGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 360
 QY 121 GluSerThrValLeuLeuAlaMetAlaPheAspArgTyrValAlaIleCysHisProLeu 140
 DB 361 GAGTCCACTGTGCTACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 420
 QY 141 ArgTyrAlaThrIleLeuThrAspThrIleIleAlaHisIleGlyValAlaValVal 160
 DB 421 CGGTATGCCACCAATCTCACTGACACCATTCATTCGCTGCTGCTGCTGCTGCTGCTG 480
 QY 161 ArgGlySerLeuLeuMetLeuProCysProPhePheIleGlyArgLeuAsnPheCysGln 180
 DB 481 CGAGGCTCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 540
 QY 181 SerHisValIleLeuHisThrTyrCysGluHisMetAlaValValIysLeuAlaCysGly 200
 DB 541 AGCATGTGATCTCACACAGTACTGTGAGCACATGCTGTGTGTGAGCTGGGCTGTGGA 600
 QY 201 AspThrArgProAsnArgValTyrGlyLeuThrAlaAlaLeuLeuValIleGlyValAsp 220
 DB 601 GACACACAGGCTTAACCGTGTGTATGGGCTGACAGCTGCATCTGTGTGTTGGGTGAC 660
 QY 221 LeuPheCysIleGlyLeuSerTyrAlaLeuIleAlaGlnAlaValLeuArgLeuSerSer 240
 DB 661 TTGTTTTCATTTGCTCTCTCTATGCTTAAAGTGCACAAAGCTGTCTTCTGCTCTCATCC 720
 QY 241 HisGluAlaArgSerIysAlaLeuGlyThrCysGlySerHisValCysValIleLeuLe 260
 DB 721 CATGAAGCTCGGTCCAAAGGCCCTAGGACCTGTGGTCCCATGCTGTGTCTCATCTCATC 780
 QY 261 SerTyrThrProAlaLeuPheSerPhePheThrHisArgPheGlyHisHisValProVal 280

781 TCTTATACACACAGCCCT 840
 281 HisIleHisIleLeuLeuAlaAsnValTyrLeuLeuProProAlaLeuAsnProVal 300
 841 CATATTCACATCTCTTTTGGCCATGTTTATCTGCTTTTGGCCACCTGCTCTTATCTCTGTG 900
 301 ValTyrGlyValIysThrIysGlnIleArgLysArgValValArgValPheGlnSerGly 320
 901 GTATATGGAGTTAAGACCAACACAGATCCGTAAAGAGATTGTCTAGGGTGTCTTCAAGTGGG 960
 321 GlnGlyMetGlyIleIysAlaSerGlu 329
 961 CAGGGAATGGGCATCAAGGCATCTGAG 987

RESULT 4
 ABK37625
 ID ABK37625 standard; cDNA; 990 BP.
 XX AC ABK37625;
 XX 08-MAY-2002 (first entry)
 XX DNA encoding G-coupled olfactory receptor #127.
 XX Human; olfactory G-coupled receptor; sensory perception of odourant;
 XX odour composition; taste composition; gene; ss.
 XX Homo sapiens.
 XX WO200198526-A2.
 XX 27-DEC-2001.
 XX 22-JUN-2001; 2001WO-US020122.
 XX 22-JUN-2000; 2000US-0213812P.
 XX 13-MAR-2001; 2001US-00804291.
 XX (SENO-) SENOMYX INC.
 XX Zozulya S, Stryer L;
 XX WPI; 2002-083330/11.
 XX P-PSDB; AAU85266.
 XX Representing sensory perception of one or more odorants for the
 XX identification and design of tastes and odors comprises providing a
 XX representative group of n olfactory receptors.
 XX Example; Page 115; 182pp; English.
 XX The invention relates to a method of representing sensory perception of
 XX one or more odorants. The method comprises: (a) providing a
 XX representative class of n olfactory receptors or ligand binding domains
 XX (LBD) of these receptors; (b) measuring values X1 to Xn representative of
 XX at least one activity of one or more odorants selected from: (i) binding
 XX one or more odorants to the LBD of at least one of the n olfactory
 XX receptors; (ii) activating at least one of the n olfactory receptors with
 XX the one or more odorants; and (iii) blocking at least one of the n
 XX olfactory receptors with the one or more odorants; and (c) generating a
 XX representation of sensory perception from the values X1 to Xn. The
 XX representation of the sensory perception of odorants is useful for the
 XX design and formulation of odour and taste compositions. ABK37499-ABK37754
 XX and ABK37918-ABK37921 represent human olfactory G-coupled receptor coding
 XX sequences and related PCR primers of the invention
 XX Sequence 990 BP; 188 A; 285 C; 221 G; 296 T; 0 U; 0 Other;

Alignment Scores:
 Pred. No.: 4,97e-169 Length: 990
 Score: 1691.00 Matches: 325
 Percent Similarity: 99.09% Conservative: 1
 Best Local Similarity: 98.78% Mismatches: 3

Query Match: 98.43% Indels: 0
 DB: 6 Gaps: 0
 US-10-081-775-2 (1-329) x ABK37625 (1-990)

QY 1 MetSerSerThrLeuGlyHisAsnMetGluSerProHisHisThrAspValAspProSer 20
 DB 1 ATGTCCAGCACTCTTGGCCACACATGGAATCTCTTAATCACACTGATGTTGACCCCTTCT 60
 QY 21 ValPhePheLeuGlyLeuProGlyLeuGluGlnPheHisLeuTrpLeuSerLeuPro 40
 DB 61 GTCCTTCTCTCTGGGCTCCAGCTCTGGACCAATTTCAITTTGGCTCTCACTCCCT 120
 QY 41 ValCysGlyLeuGlyThrAlaThrLeuValGlyAsnLeuThrLeuValValValAla 60
 DB 121 GGTGTGGCTTAGGCACAGCCACATTTGGGCAATATAACTATTCTGGTTTGTGGTCC 180
 QY 61 ThrGluProValLeuHisLysProValTyrLeuPheLeuCysMetLeuSerThrIleAsp 80
 DB 181 ACTGAACCAAGCTTGGCAAGCCTGTGTACCTTTTCTGTGATGCTCTCAACCATCGAC 240
 QY 81 LeuAlaAlaSerValSerThrValProLysLeuLeuAlaIlePheTrpCysGlyAlaGly 100
 DB 241 TTGGCTGCTCTGTCTCCACAGTCTCCCAAGCTACTGGCTATCTTCTGATGCTCTCAACCATCGAC 300
 QY 101 HisIleSerAlaSerAlaCysLeuAlaHisMetPhePheIleHisAlaPheCysMetMet 120
 DB 301 CATATATCTGCTCTGCTGCTGGCAGCAGATGTTCTTCAITTCATGCTCTTGCATGATG 360
 QY 121 GluSerThrValLeuLeuAlaMetAlaPheAspArgTyrValAlaIleCysHisProLeu 140
 DB 361 GAGTCCACATGTGTACTGGCCATGGCTTTGATGCTAGTGGCCATCTGCCACCCATC 420
 QY 141 ArgTyrAlaThrIleLeuThrAspThrIleIleAlaHisIleGlyValAlaValVal 160
 DB 421 CGCTATGCCAATCTCTCAGTACACCATCATATGCCACATAGGGTGGCAGCTGTAGT 480
 QY 161 ArgGlySerLeuLeuMetLeuProCysProPhePheIleGlyArgLeuAsnPhCysGln 180
 DB 481 CGAGGCTCCCTGCTCATGCTCCCATGTCCTTCTTATGTCGCTTGAATCTTGCCTA 540
 QY 181 SerHisValIleLeuHisThrTyrCysGluHisMetAlaValValLysLeuAlaCysGly 200
 DB 541 AGCCATGTGATCTCTACACACATGCTGTGAGCAGATGCTGTGGTGAAGCTGGCTGTGGA 600
 QY 201 AspThrArgProAsnArgValTyrGlyLeuThrAlaAlaLeuLeuValIleGlyValAsp 220
 DB 601 GACACCAAGGCTTAACCGTGTGTATGGCTGACAGCTGACCTGTGTGTGTTGGGTTGAC 660
 QY 221 LeuPheCysIleGlyLeuSerTyrAlaLeuIleAlaGlnAlaValLeuArgLeuSerSer 240
 DB 661 TTGTTTTCATTTGGTCTCTCTATGCCCTAAGTGCAAGCTGTGCTTCGCCCTCTCATCC 720
 QY 241 HisGluAlaArgSerLysAlaLeuGlyThrCysGlySerHisValCysValIleLeuIle 260
 DB 721 CATGAAGCTCGTGCCAAAGGCTTAGGACCTGTGTTTCCCATGCTGTGTCTATCTTCATC 780
 QY 261 SerTyrThrProAlaLeuPheSerPhePheThrHisArgPheGlyHisHisValProVal 280
 DB 781 TCTTATACACAGCCCTCTCTCTATGCCCTAAGTGCAAGCTGTGCTTCGCCCTCTCATCC 840
 QY 281 HisIleHisIleLeuLeuAlaAsnValTyrLeuLeuLeuProAlaLeuAsnProVal 300
 DB 841 CATATTACATCTTTTGGCCATGTTTATCTGCTTTTGGCCATCTCTTAATCTCTG 900
 QY 301 ValTyrGlyValLysThrLysGlnIleArgLysArgValValArgValPheGlnSerGly 320
 DB 901 GTATATGGAGTTAAGCAACAGATCCGTAAGAGTTGTTCAGGGTGTTCAGAGTGGG 960
 QY 321 GlnGlyMetGlyIleLysAlaSerGlu 329
 DB 961 CAGGGAATGGCATCAAGGCATCTGAG 987

RESULT 5
 ABQ88371
 ID ABQ88371 standard; cDNA; 998 BP.
 XX
 AC ABQ88371;
 XX
 DT 23-SEP-2002 (first entry)
 XX
 DE Human G protein coupled receptor cDNA SEQ ID NO:35.
 XX
 KW Human; G protein coupled receptor; GPCR; GPCRX; neuroprotective;
 KW neotropic; anti-HIV; antiasthmatic; antiarteriosclerotic; cyostatic;
 KW immunomodulator; antiinflammatory; antidiabetic; anorectic; haemostatic;
 KW antibacterial; fungicide; protozoal; virucide; nephrotropic; osteopathic;
 KW cardiant; antulcer; antiallergic; hepatotropic; antiparkinsonian; HIV;
 KW vaccine; gene therapy; cell signal processing; cardiomyopathy; diabetes;
 KW metabolic pathway modulation; atherosclerosis; cancer; obesity; asthma;
 KW infection; Parkinson's disease; osteoporosis; Crohn's disease; ulcer;
 KW allergy; cirrhosis; glomerulonephritis; stroke; haematopoietic disorder;
 KW systemic lupus erythematosus; gene; ss.
 XX
 OS Homo sapiens.
 XX
 PN WO200250276-A2.
 XX
 PD 27-JUN-2002.
 XX
 PF 18-DEC-2001; 2001WO-US049347.
 XX
 PR 18-DEC-2000; 2000US-0256635P.
 PR 21-DEC-2000; 2000US-0257876P.
 PR 04-JAN-2001; 2001US-0259743P.
 PR 10-JAN-2001; 2001US-0260718P.
 PR 12-JAN-2001; 2001US-0261498P.
 PR 24-JAN-2001; 2001US-0263689P.
 PR 08-FEB-2001; 2001US-0267454P.
 PR 22-FEB-2001; 2001US-0271021P.
 PR 14-MAR-2001; 2001US-0275946P.
 PR 23-MAR-2001; 2001US-0278150P.
 PR 18-APR-2001; 2001US-0284591P.
 PR 23-APR-2001; 2001US-0285718P.
 PR 19-JUN-2001; 2001US-0299327P.
 PR 16-AUG-2001; 2001US-0312902P.
 XX
 (CURA-) CURAGEN CORP.
 PA
 XX
 XX
 PI Li L, Padigaru M, Ballinger RA, Kekuda R, Colman SD, Sciore P;
 PI Smithson G, Peyman JA, Macdougall JR, Stone D, Vernet CAM, Shenoy S;
 PI Gunther E, Millet I, Tchernev VT, Anderson D, Gusev V, Malyankar UM;
 PI Zhong H, Ellerman KE, Wolenc A;
 XX
 XX
 WIPI; 2002-557660/59.
 P-PSDB; ABP51577.
 XX
 PT New isolated human G-protein coupled receptor X (GPCRX) polypeptide,
 PT useful for treating or preventing GPCR-associated disorders e.g.
 PT diabetes, atherosclerosis, cancer or obesity.
 XX
 PS Claim 8; Page 88; 354pp; English.
 XX
 CC ABQ88354 to ABQ88417 represent human G protein coupled receptor (GPCR)
 CC cDNA sequences, and ABP51560 to ABP51624 represent human GPCR proteins
 CC from the present invention. GPCR sequences can have neuroprotective,
 CC neotropic, anti-HIV, antiasthmatic, antiarteriosclerotic, cyostatic,
 CC immunomodulator, antiinflammatory, antidiabetic, anorectic, haemostatic,
 CC antibacterial, fungicide, protozoal, virucide, nephrotropic, osteopathic,
 CC cardiant, antulcer, antiallergic, hepatotropic and antiparkinsonian
 CC activities, and can be used in vaccines and gene therapy. GPCR proteins,
 CC nucleic acid molecules, and antibodies from the present invention can be
 CC used for manufacturing a medicament for treating or preventing a GPCR-
 CC associated disorder or syndrome related to cell signal processing and
 CC metabolic pathway modulation, such as cardiomyopathy, atherosclerosis,
 CC diabetes, cancer, obesity, infections (bacterial, fungal, protozoal or

US-10-081-775-2 (1-329) X ABO88371 (1-998)

Db	782	TCCTATACACACCCCTCTCTCTCTCTCTCTCTTTTACACACCGCTTTGGCCATCAGTTCAGTC	841		
Qy	281	HisIleHisIleLeuLeuAlaAsnValTyrLeuLeuLeuProAlaLeuAsnProVal	300		
Db	842	CAATTCACATCTTTTGGCAATGTTATCTGCTTTTGCACCTGCTTTAATCTCTGTG	901		
Qy	301	ValTyrGlyValLysThrLysGlnLeuArgLysArgValValArgValPheGlnSerGly	320		
Db	902	GTATATGGAGCTTAAGACCAACACAGATCCGTAAAGAGTTCCTCAGGGTGTTCACAAAGTGGG	961		
Qy	321	GlnGlyMetGlyIleLysAlaSerGlu	329		
Db	962	CAGGGAATGGGCATCAAGGCATCTGAG	988		
RESULT 6					
ID	ABQ88372	standard; cDNA; 998 BP.			
XX	ABQ88372;				
XX	23-SEP-2002	(first entry)			
DE	Human G protein coupled receptor cDNA SEQ ID NO:37.				
XX	Human; G protein coupled receptor; GPCR; GPCR; neuroprotective;				
XX	nootropic; anti-HIV; antiasthmatic; antiarteriosclerotic; cytostatic;				
KW	immunomodulator; antiinflammatory; antidiabetic; anorectic; haemostatic;				
KW	antibacterial; fungicide; protozoal; virucide; nephrotropic; osteopathic;				
KW	cardiant; antiulcer; antiallergic; hepatotropic; antiparkinsonian; HIV;				
KW	vaccine; gene therapy; cell signal processing; cardiomyopathy; diabetes;				
KW	metabolic pathway modulation; atherosclerosis; cancer; obesity; asthma;				
KW	infection; Parkinson's disease; osteoporosis; Crohn's disease; ulcer;				
KW	allergy; cirrhosis; glomerulonephritis; stroke; haematopoietic disorder;				
KW	systemic lupus erythematosus; gene; ss.				
XX	Homo sapiens.				
OS					
XX	WO200250276-A2.				
XX	27-JUN-2002.				
XX	18-DEC-2001;	2001WO-US049347.			
XX	18-DEC-2000;	2000US-0256635P.			
PR	21-DEC-2000;	2000US-0257873P.			
PR	04-JAN-2001;	2001US-0259743P.			
PR	10-JAN-2001;	2001US-0260718P.			
PR	12-JAN-2001;	2001US-0261498P.			
PR	24-JAN-2001;	2001US-0263689P.			
PR	08-FEB-2001;	2001US-0267464P.			
PR	22-FEB-2001;	2001US-0271021P.			
PR	14-MAR-2001;	2001US-0275946P.			
PR	23-MAR-2001;	2001US-0278150P.			
PR	18-APR-2001;	2001US-0284591P.			
PR	23-APR-2001;	2001US-0285718P.			
PR	19-JUN-2001;	2001US-0299327P.			
PR	16-AUG-2001;	2001US-0312902P.			
XX	(CURA-) CURAGEN CORP.				
XX	Li L, Padigar M, Ballinger RA, Kekuda R, Colman SD, Sciore P;				
PI	Smithson G, Feyman JA, Macdougall JR, Stone D, Vernet CAM, Shenoy SI;				
PI	Gunther E, Millet I, Tchernev VT, Anderson D, Gusev V, Malyankar UN;				
PI	Zhong H, Ellerman KE, Wolenc A;				
XX	WPI; 2002-557660/59.				
DR	P-PSDB; ABP51578.				
XX	New isolated human G-protein coupled receptor X (GPCRX) polypeptide,				
PT	useful for treating or preventing GPCR-associated disorders e.g.				
PT	diabetes, atherosclerosis, cancer or obesity.				
XX					

PS Claim 8; Page 88; 354pp; English.

CC ABQ88354 to ABQ88417 represent human G protein coupled receptor (GPCR) cDNA sequences, and ABP51560 to ABP51624 represent human GPCR proteins from the present invention. GPCR sequences can have neuroprotective, CC nootropic, anti-HIV, antiasthmatic, antiarteriosclerotic, cytostatic, CC immunomodulator, antiinflammatory, antidiabetic, anorectic, haemostatic, CC antibacterial, fungicide, protozoal, virucide, nephrotropic, osteopathic, CC cardiant, antiulcer, antiallergic, hepatotropic and antiparkinsonian CC activities, and can be used in vaccines and gene therapy. GPCR proteins, CC nucleic acid molecules, and antibodies from the present invention can be CC used for manufacturing a medicament for treating or preventing a GPCR- CC associated disorder or syndrome related to cell signal processing and CC metabolic pathway modulation, such as cardiomyopathy, atherosclerosis, CC diabetes, cancer, obesity, infections (bacterial, fungal, protozoal or CC viral), HIV, asthma, Parkinson's disease, osteoporosis, Crohn's disease, CC ulcers, allergies, cirrhosis, glomerulonephritis, stroke, systemic lupus CC erythematosus, or haematopoietic disorders. Anti-GPCR antibodies can be CC used diagnostically to monitor protein levels in tissues as part of a CC clinical testing procedure such as in determining the efficacy of a given CC treatment regimen. ABQ88418 to ABQ88639 represent PCR primers and probes CC for the human GPCRs of the present invention

XX Sequence 998 BP; 190 A; 287 C; 223 G; 298 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.: 5.03e-169 Length: 998
Score: 1691.00 Matches: 325
Percent Similarity: 99.09% Conservative: 1
Best Local Similarity: 98.78% Mismatches: 3
Query Match: 98.43% Indels: 0
DB: 6 Gaps: 0

US-10-081-775-2 (1-329) x ABQ88372 (1-998)

Qy 1 MetSerThrLeuGlyHisAsnMetGluSerProHisHisThrAspValAspProSer 20
Db 2 ATGTCCAGCACTCTGTGGCCAAACATGGAATCTCTTAATCACTGATGTTGACCCCTCT 61
Qy 21 ValPhePheLeuLeuGlyIleProGlyLeuGluGlnPheHisLeuTyrLeuSerLeuPro 40
Db 62 GTCTTCTCTCTGCGCATCCCAAGCTGTGGCAATATTAATACTTCTGGTGTGTGGCC 121
Qy 41 ValCysGlyLeuGlyThrAlaThrIleValGlyAsnIleThrIleLeuValValAla 60
Db 122 GTGTGTGGCTTAGGCACAGCCACAATGTGGCAATATAATACTTCTGGTGTGTGGCC 181
Qy 61 ThrGluProValLeuHisLysProValTyrIleuPheLeuCysMetLeuSerThrIleAsp 80
Db 182 ACTGAACCACTGTGCACAGCCCTGTGTACCTTTTCTGTGCATGCTCTCAACCATCGAC 241
Qy 81 LeuAlaAlaSerValSerThrValProLysLeuLeuAlaIlePheTyrCysGlyAlaGly 100
Db 242 TTGGCTGCCTCTGTCTCCACAGTTCACAGCTACTGGCTATCTTCTGGTGTGGAGCCGA 301
Qy 101 HisIleSerAlaSerAlaCysLeuAlaHisMetPhePheIleHisAlaPheCysMetMet 120
Db 302 CATATATCTGCCTCTGCCTGCGCTGGCAGACATGTTCTTCAATCATGCCCTTCTGCATGATG 361
Qy 121 GluSerThrValLeuLeuAlaMetAlaPheAspArgTyrValAlaIleCysHisProLeu 140
Db 362 GAGTCCACTGTGTACTGGCCATGGCTTGTATGCTGCTAGCTGCGCCATCTGCCACCATC 421
Qy 141 ArgTyrAlaThrIleLeuThrAspThrIleIleAlaHisIleGlyValAlaIleVal 160
Db 422 CGTATATGCCACAATCTCTACAGACATCATTTGCCCATATGGCCATAGGGGTGGCAGCTGATG 481
Qy 161 ArgGlySerLeuLeuMetLeuProCysProPhePheIleGlyArgLeuAsnPheCysGln 180
Db 482 CGAGGCTCCCTGCTATGCTCCATGCTCCATGCTCTTCTTATTTGGCGTTTGAACCTTCTGCCAA 541
Qy 181 SerHisValIleLeuHisThrTyrCysGluHisMetAlaValValLysLeuAlaCysGly 200

Db 542 AGCCATGTGATCTCTACACAGTACTGTGAGCACATGCTGTGCTGAAGCTGGCTGTGGA 601
Qy 201 AspThrArgProAsnArgValTyrGlyLeuThrAlaAlaLeuLeuValIleGlyValAsp 220
Db 602 GACACAGGCGCTAACCGTGTGTATGGCTGACAGCTGCATGTGTGTCTATGGGTGAC 661
Qy 221 LeuPheCysIleGlyLeuSerTyrAlaLeuIleAlaGlnAlaValLeuArgLeuSerSer 240
Db 662 TTGTTTGTGATGGTCTCTCTCTATGCCCTAAGTGGACAGAGCTGCTTGGCCTCTCATCC 721
Qy 241 HisGluAlaArgSerLysAlaLeuGlyThrCysGlySerHisValCysValIleLeuIle 260
Db 722 CATGAAGCTCGGTGCAAGGCCCTAGGACCTGTGGTTCCCATGCTCTGTGTCTCATCTCATC 781
Qy 261 SerTyrThrProAlaLeuPheSerPhePheThrHisArgPheGlyHisHisValProVal 280
Db 782 TCTTATACACAGCCCTCTTCTCTCTTTTACACACCGCTTTGGCCATCACGTTCCAGTC 841
Qy 281 HisIleHisIleLeuLeuAlaAsnValTyrLeuLeuLeuProAlaLeuAsnProVal 300
Db 842 CATATTCAATTCTTTTGGCCAAATGTTTATCTGTCTTTGGCACCTGCTTTAATCTGTG 901
Qy 301 ValTyrGlyValLysThrLysGlnIleArgLysArgValValArgValPheGlnSerGly 320
Db 902 GTATATGGAGTTAAGACCAACAGATCCGTAAACAGATTGTCAAGGCTGTTTCAAAAGTGG 961
Qy 321 GlnGlyMetGlyIleLysAlaSerGlu 329
Db 962 CAGGGAATGGGCATCAAGGCATCTGAG 988
RESULT 7
ADC86344
ID ADC86344 standard; DNA; 1390 BP.
XX AC ADC86344;
XX DT 01-JAN-2004 (first entry)
XX DE Human GPCR gene SEQ ID NO:797.
XX KW ds; gene; human; GPCR;
XX OS Guanosine triphosphate-binding protein coupled receptor; gene therapy.
XX FN Homo sapiens.
XX EP1270724-A2.
XX PD 02-JAN-2003.
XX PF 18-JUN-2002; 2002EP-00013517.
XX PR 18-JUN-2001; 2001JP-00245789.
XX PA (NAAD-) NAT INST ADVANCED IND SCI & TECHNOLOGY.
XX PA (ADSC-) CENT ADVANCED SCI & TECHNOLOGY INCUBATIO.
XX PI Suwa M, Asai K, Akiyama Y, Aburatani H;
XX WPI; 2003-315783/31.
XX P-PSDB; ADC86345.
XX PT New polynucleotide, useful for preparing a composition for treating a
XX PT patient in need of increased or suppressed activity or expression of the
XX XX guanosine triphosphate-binding protein coupled receptor.
XX PS Claim 1; SEQ ID NO 797; 28pp; English.
XX CC The invention relates to a novel polynucleotide encoding a guanosine
XX CC triphosphate-binding protein coupled receptor (GPCR). A polynucleotide of
XX CC the invention may have a use in gene therapy. The polynucleotide and
XX CC polypeptide are useful for preparing a composition for treating a patient
XX CC in need of increased or suppressed activity or expression of the
XX CC guanosine triphosphate-binding protein coupled receptor. The

CC polynucleotide sequences shown in ADC85548-ADC87616 encode GPCR's of the
 CC invention.
 XX
 SQ Sequence 1390 BP; 317 A; 353 C; 293 G; 427 T; 0 U; 0 Other;

Alignment Scores:
 Pred. No.: 8,03e-169 Length: 1390
 Score: 1691.00 Matches: 325
 Percent Similarity: 99.09% Conservative: 1
 Best Local Similarity: 98.78% Mismatches: 3
 Query Match: 98.43% Indels: 0
 DB: 9 Gaps: 0

US-10-081-775-2 (1-329) x ADC86344 (1-1390)
 QY 1 MetSerThrLeuGlyHisAsnMetGluSerProHisHisThrAspValAspProSer 20
 Db 201 ATGTCCAGCACTCTTGGCCACCAATGGAATCTCTTAATCACACTGATGTTGACCTTCT 260
 QY 21 ValPheLeuLeuGlyIleProGlyLeuGluGlnPheHisLeuTrpLeuSerLeuPro 40
 Db 261 GTCTTCTCTCCCTGGGCATCCAGGCTGGACAAATTTCAATTTGTGGCTCTCACTCCCT 320
 QY 41 ValCysGlyLeuGlyThrAlaThrIleValGlyAsnIleThrIleLeuValValAla 60
 Db 321 GTGTGTGGCTTAGGCACAGCCACCAATTTGGGCAATATTAATCTCTGTTGTGTGCC 380
 QY 61 ThrGluProValLeuHisLeuProValTrpLeuPheLeuCysMetLeuSerThrIleAsp 80
 Db 381 ACTGACCAAGCTTGGCACAAGCTGTGTACCTTTTCTGTGCATGCTCTCAACCATCGAC 440
 QY 81 LeuAlaAlaSerValSerThrValProLysLeuLeuAlaIlePheTrpCysGlyValagly 100
 Db 441 TTGGCTGCTCTGTCTCCACAGTTCCCAAGCTACTGGCTATCTTCTGTGTGGAGCCGGA 500
 QY 101 HisIleSerAlaSerAlaCysLeuAlaHisMetPhePheIleHisAlaPheCysMetMet 120
 Db 501 CATATATCTGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 560
 QY 121 GluSerThrValLeuLeuAlaMetAlaPheAspArgTrpValAlaIleCysHisProLeu 140
 Db 561 GAGTCCACTGTGCTACTGGCCATGGCTTTGATGCTAGCTGGCCATCTGCCACCACTC 620
 QY 141 ArgTrpAlaThrIleLeuLeuThrAspThrIleIleAlaHisIleGlyValAlaValVal 160
 Db 621 CGCTATGCCAATCTCACTGACACCATCATTTGCCACATAGGGGTGGCAGCTGTAGTG 680
 QY 161 ArgGlySerLeuLeuMetLeuProCysProPhePheIleGlyValArgLeuAsnPheCysGln 180
 Db 681 CGAGGCTCCCTGCTCATGCTCCATGCTCCATGCTCCATGCTCCATGCTCCATGCTCCATG 740
 QY 181 SerHisValIleLeuHisThrTrpCysGluHisMetAlaValValLysLeuAlaCysGly 200
 Db 741 AGCCATGTGATCTACACAGTACTGTGAGCAGCATGGCTGTGTGTAAGCTGGCCTGTGA 800
 QY 201 AspThrArgProAsnArgValTrpGlyLeuThrAlaAlaLeuLeuValIleGlyValAsp 220
 Db 801 GACACCAAGCCTAACCGTGTGTATGGGCTGACAGCTGCACTGTGTGTGTGTGTGTGTGTG 860
 QY 221 LeuPheCysIleGlyLeuSerTrpAlaLeuIleAlaGlnAlaValLeuArgLeuSerSer 240
 Db 861 TTGTTTGTGATTTGTTCTCTCTATGCTTAAAGTGCACAGTGTCTCTGCTCTCATCC 920
 QY 241 HisGluAlaArgSerLysAlaLeuGlyThrCysGlySerHisValCysValIleLeuIle 260
 Db 921 CATGAGCTCGTCCCAAGGCCCTTAGGACCTGTGGTGTCCATGCTGTGTGTGTGTGTGTGTG 980
 QY 261 SerTrpThrProAlaLeuPheSerPhePheThrHisArgPheGlyHisHisValProVal 280
 Db 981 TCTTATACACAGCCCT 1040
 QY 281 HisIleHisIleLeuLeuAlaAsnValTrpLeuLeuLeuProAlaLeuAsnProVal 300

Db 1041 CATATTACATCTTTTGGCCAAATTTATCTCTTTGGCCACCTCTCTTATCTCTGTG 1100
 QY 301 ValTyrGlyValLysThrLysGlnIleArgLysArgValValArgValPheGlnSerGly 320
 Db 1101 GTATATGGAGTTTAAGACCAACAGATCCGTAAAGAGTTGTCAAGGTGTTTCAAGTGGG 1160
 QY 321 GlnGlyMetGlyIleLysAlaSerGlu 329
 Db 1161 CAGGGAATGGCATCAAGGCATCTGAG 1187

RESULT 8
 AAH31850
 ID AAH31850 standard; DNA; 963 BP.

XX AAH31850;
 AC
 XX 30-JUL-2001 (first entry)
 DT
 XX
 XX Human olfactory receptor polynucleotide, SEQ ID NO: 423.

XX Human; olfactory receptor; OR; primary scent determination;
 KW secondary scent determination; polypeptide library; odour receptor;
 KW scent profile; scent fingerprint; scent representation; ds.

XX Homo sapiens.

XX WO200127158-A2.

XX 19-APR-2001.

XX 06-OCT-2000; 2000MO-US027582.

XX 08-OCT-1999; 99US-0158615P.

XX 24-FEB-2000; 2000US-0184809P.

XX (DIGI-) DIGISCENTS.

XX (YEDA) YEDA RES & DEV CO LTD.

XX Bellenston J, Smith D, Lancet D, Glusman G, Fuchs T, Yanai I;

XX WPI; 2001-290713/30.

XX New polynucleotides which encode polypeptides involved in olfactory
 sensation for identifying olfactory agonists and antagonists.

XX Claim 8; Page 349; 1857pp; English.

XX The present sequence is one of a number of isolated polynucleotides which
 encode polypeptides involved in olfactory sensation. The polynucleotides
 can be used in screening for olfactory agonists and antagonists. The
 methods allow for the determination of primary scents and the
 identification of the odour receptors used to detect these primary
 scents. The methods also enable determination of secondary scents and the
 identification of combinations of odour receptors that are involved in
 detecting such secondary scents. This enables the construction of a scent
 representation (also called a scent fingerprint or scent profile), which
 may be used to re-create and edit scents. Libraries of olfactory
 receptors are useful for determining the interaction pattern of a
 composition with the receptors, and can be used for determining
 differences in the olfactory faculties of different individuals

XX Sequence 963 BP; 181 A; 276 C; 216 G; 290 T; 0 U; 0 Other;

Alignment Scores:
 Pred. No.: 1,33e-164 Length: 963
 Score: 1649.00 Matches: 317
 Percent Similarity: 99.07% Conservative: 1
 Best Local Similarity: 98.75% Mismatches: 3
 Query Match: 95.98% Indels: 0
 DB: 4 Gaps: 0

US-10-081-775-2 (1-329) x AAH31850 (1-963)


```

Alignment Scores:
Pred. No.:      1,34e-164      Length:      966
Score:          1849.00      Matches:     317
Percent Similarity: 99.07%      Misconservative: 1
Best Local Similarity: 98.75%      Mismatches:    3
Indels:         95.98%      Gaps:         0
Query Match:    6
DB:

US-10-081-775-2 (1-329) x ABK68612 (1-966)

Qy  9  MetGluSerProHisHisThrAspValAppProSerValPhePheLeuGlyIlePro 28
      |||:::|||||
Db   1  ATGGAATCTCCTAATCACACTGATGTGACCCCTCTGTCTTCTCTCTCTGGGCATCCCA 60

```

RESULT 10
ABK68612
ID ABK68
XX
AC ABK68
XX
DT 02-JUN
XX

Human G protein coupled receptor coding sequence SEQ ID 205.

Human; gene; anorectic; antiasthmatic; antidiabetic; hypotensive; antiparkinsonian; nootropic; neuroprotective; tranquiliser; antirheumatic; antiinflammatory; osteopathic; cardant; neuroleptic; antiarthritic; gene therapy; olfactory G protein-coupled receptor; GPCR; infection; obesity; diabetes; hypertension; malnutrition; Parkinson's disease; Alzheimer's disease; Korsakoff's psychosis; anxiety; rheumatoid arthritis; chronic obstructive pulmonary disease; osteoporosis; asthma; myocardial infarction; schizophrenia; osteoarthritis; ds.

Homo sapiens.

WO2003000735-A2.

03-JAN-2003.

24-JUN-2002; 2002WO-IB002481.

26-JUN-2001; 2001US-030109SP.

06-NOV-2001; 2001US-0332759P.

(DECO-) DECODE GENETICS EHF.

Martinez RAM, Sigurdsson GT; WPI; 2003-175284/17.

P-PSDB; ABR01673.

New olfactory G protein-coupled receptor gene nucleic acid and polypeptide, useful for diagnosing or treating a disease or condition associated with GPCR, e.g. obesity, diabetes, hypertension, malnutrition or Alzheimer's disease.

Claim 1; Page 103; 383pp; English.

The present invention relates to novel human olfactory G protein-coupled receptors (GPCR) and their coding sequences (ABZ77872-ABZ77986 and ABR01571-ABR01685). The GPCRs and coding sequences are useful for diagnosing or treating a disease or condition associated with GPCR, e.g. infections, obesity, diabetes, hypertension, malnutrition, Parkinson's disease, Alzheimer's disease, Korsakoff's psychosis, anxiety, rheumatoid arthritis, chronic obstructive pulmonary disease, osteoporosis, asthma, myocardial infarction, schizophrenia, or osteoarthritis

Sequence 948 BP; 174 A; 266 C; 202 G; 306 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.:	1.01e-135	Length:	948
Score:	1376.00	Matches:	267
Percent Similarity:	94.16%	Conservative:	7
Best Local Similarity:	91.75%	Mismatches:	17
Query Match:	80.09%	Indels:	0
DB:	7	Gaps:	0

US-10-081-775-2 (1-329) x ABZ77974 (1-948)

Qy 20 SerValPhePheLeuLeuGlyIleProGlyLeuGluGlnPheHisLeuTrpLeuSerLeu 39
Db 28 TCTGAGTTATCTGATGGATTCCTGGCATTCACACTTCTGTTTGTGTTTTTTT 87
Qy 40 ProValCysGlyLeuGlyThrAlaThrIleValGlyAsnIleThrIleLeuValVal 59
Db 88 TTGTTGTTGTTGTTGTTGTTTTCACAAATGTTGGCAATAAATACTATTCGGTGTGTT 147
Qy 60 AlaThrGluProValLeuHisIysProValTyrLeuPheLeuCysMetLeuSerThrIle 79
Db 148 GCACACTGAACCACTCTGCACAGCTTCACAGTATCCCAAGCTATCTGCTGTGGAGCC 207
Qy 80 AspLeuAlaAlaSerValSerThrValProLysLeuLeuAlaIlePheTrpCysGlyAla 99
Db 208 GACTTGGCTGCTGCTGCTTCACAGTATCCCAAGCTATCTGCTGTGGAGCC 267

The present invention relates to novel human olfactory G protein-coupled receptors (GPCR) and their coding sequences (AB277872-AB277886 and AB277887-AB277895). The GPCRs and coding sequences are useful for diagnosing or treating a disease or condition associated with GPCR, e.g. infections, obesity, diabetes, hypertension, malnutrition, Parkinson's disease, Alzheimer's disease, Korsakoff's psychosis, anxiety, rheumatoid arthritis, chronic obstructive pulmonary disease, osteoporosis, asthma.

ms Page Book (uspio)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 27, 2004, 14:41:51 ; Search time 478 Seconds

(without alignments)
8771.907 Million cell updates/sec

Title: US-10-081-775-1_COPY_537_1523

Perfect score: 987

Sequence: 1 atgtccagcactcttggcca.....tgggcatcaaggcatcttgag 987

Scoring table: IDENTITY_NUC

; Gapop 10.0 , Gapext 1.0

Searched: 3373863 seqs, 2124099041 residues

Total number of hits satisfying chosen parameters: 6747725

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : N_Geneseq_29Jan04:*

- 1: Geneseq1980s:*
- 2: Geneseq1990s:*
- 3: Geneseq2000s:*
- 4: Geneseq2001as:*
- 5: Geneseq2001bs:*
- 6: Geneseq2002s:*
- 7: Geneseq2003as:*
- 8: Geneseq2003bs:*
- 9: Geneseq2003cs:*
- 10: Geneseq2004s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	987	100.0	1567	8	ACD91434 Human G-p
2	980.6	99.4	990	5	AAS42339 Human cdn
3	980.6	99.4	990	6	ABZ43046 Human GPC
4	980.6	99.4	990	6	ABK37625 DNA encod
5	980.6	99.4	998	6	ABQ88371 Human G p
6	980.6	99.4	998	6	ABQ88372 Human G p
7	980.6	99.4	1390	9	ADC86344 Human GPC
8	956.6	96.9	963	4	Aah31850 Human olf
9	956.6	96.9	966	6	ABK16633 Human G-c
10	956.6	96.9	966	6	ABK68612 Human G-c
11	791	80.1	948	7	ABZ77974 Human DNA
12	376.2	38.1	980	4	AAD19143 Human G p
13	376.2	38.1	1986	4	AAD113897 Human G-p
14	376.2	38.1	1986	4	ABA55607 Human G-p
15	376.2	38.1	1986	4	ABA55607 Human G-p
16	376.2	38.1	1986	4	AAI35262 Human G-p
17	376.2	38.1	1986	4	ABA45127 Human G-p
18	376.2	38.1	1986	4	ABA25305 Human G-p
19	376.2	38.1	1986	4	AAK29302 Human G-p
20	376.2	38.1	1986	4	AAK03835 Human G-p
21	376.2	38.1	1986	4	ABZ8924 Human G-p
22	376.2	38.1	1986	5	AAI03764 Human G-p
23	376.2	38.1	1986	6	ABZ03861 Human G-p
24	375.2	38.0	942	4	AAH32015 Human G-p

24	375.2	38.0	945	4	AAD12947	Aad12947 Human G-p
25	375.2	38.0	945	5	AAS42253	Aas42253 Human cdn
26	375.2	38.0	945	6	AAL44139	Aal44139 Human G-p
27	375.2	38.0	945	6	ABZ42950	Abz42950 Human GPC
28	375.2	38.0	945	6	ABK37539	Abk37539 DNA encod
29	374.6	38.0	1345	9	ADC86304	Adc86304 Human GPC
30	373.6	37.9	945	6	ABK68589	Abk68589 Human DNA
31	369.8	37.5	945	4	AAI23114	Aai23114 Probe #13
32	369.8	37.5	945	4	ABA68205	AbA68205 Human G-p
33	369.8	37.5	945	4	AAI48420	Aai48420 Probe #17
34	369.8	37.5	945	4	ABA50263	AbA50263 Human G-p
35	369.8	37.5	945	4	ABA35217	AbA35217 Probe #13
36	369.8	37.5	945	4	AAK42341	Aak42341 Human G-p
37	369.8	37.5	945	4	AAK16586	Aak16586 Human G-p
38	369.8	37.5	945	4	ABSA1952	AbsA1952 Human G-p
39	369.8	37.5	945	5	AAI08767	Aai08767 Probe #87
40	369.8	37.5	945	6	ABS16397	AbS16397 Human G-p
41	368.4	37.3	1013	6	ABQ88366	Abq88366 Human G-p
42	368.4	37.3	1360	9	ADC86368	Adc86368 Human GPC
43	367.2	37.2	982	4	AAD19141	Aad19141 Human G-p
44	365	37.0	1345	9	ADC86302	Adc86302 Human GPC
45	364	36.9	942	4	AAH32013	Aah32013 Human G-p

ALIGNMENTS

RESULT 1

ACD91434	ACD91434 standard; cdna; 1567 BP.
XX	XX
AC	ACD91434;
XX	XX
DT	22-SEP-2003 (first entry)
XX	XX
DE	Human G-protein coupled receptor HGRPRMY25 cdna.
XX	XX
KW	Human; G-protein coupled receptor; HGRPRMY25; immune disease;
KW	inflammatory disease; arthritis; asthma; AIDS; psoriasis;
KW	graft-versus-host disease; systemic lupus erythematosus;
KW	reproductive disorder; varicocele; orchitis; neural disorder;
KW	Alzheimer's disease; Parkinson's disease; depression; schizophrenia;
KW	cardiovascular disorder; hypertension; acute heart failure;
KW	pulmonary disorder; endocrine disorder; obesity; diabetes; anorexia;
KW	bone disorder; osteoporosis; pain; cancer; chromosome identification;
KW	gene therapy; gene; ss.
OS	Homo sapiens.
XX	XX
PN	US2003060409-A1.
XX	XX
PD	27-MAR-2003.
XX	XX
PF	21-FEB-2002; 2002US-00081775.
XX	XX
PR	21-FEB-2001; 2001US-0270134P.
XX	XX
XX	27-MAR-2001; 2001US-0278952P.
XX	XX
PA	(RAMA/) RAMANATHAN C S.
XX	XX
PA	(FEDE/) FEDER J N.
XX	XX
PA	(MINT/) MINTIER G A.
XX	XX
PI	Ramanathan CS, Feder JN, Mintier GA;
XX	XX
DR	WPI; 2003-521919/49.
XX	XX
DR	P-PSDB; ABO42809.
XX	XX
XX	XX
PT	New nucleic acid molecule encoding a human G-protein coupled receptor
XX	XX
PT	(HGRPRMY25) is useful for diagnosing, preventing or treating diseases
XX	XX
PT	involving the receptor, e.g. inflammation, diabetes, asthma, hypertension
XX	XX
XX	XX
PS	Claim 1; Fig 1A-B; 139pp; English.

XX The invention describes an isolated nucleic acid molecule comprising a
CC sequence that is at least 95% identical to a polynucleotide encoding
CC novel human G-protein coupled receptor HGPBMV25. The nucleic acid
CC molecule, polypeptide and antibody are useful in diagnosing, preventing,
CC treating or ameliorating medical conditions where GPCR is directly or
CC indirectly involved, such as immune or inflammatory diseases (e.g.
CC arthritis, asthma, AIDS, graft-versus-host disease, psoriasis or systemic
CC lupus erythematosus), reproductive disorders (e.g. varicocele or
CC orchitis), neural disorders (e.g. Alzheimer's disease, Parkinson's
CC disease, depression or schizophrenia), cardiovascular disorders (e.g.
CC hypertension or acute heart failure), pulmonary disorders, endocrine
CC disorders (e.g. obesity, diabetes or anorexia), bone disorders (e.g.
CC osteoporosis), pain or cancer. The polynucleotide may also be used in
CC chromosome identification, in identifying organisms from minute
CC biological samples, or as molecular weight markers. This sequence encodes
CC a novel human G-protein coupled receptor HGPBMV25
XX
SQ Sequence 1567 BP; 367 A; 382 C; 314 G; 504 T; 0 U; 0 Other;
Query Match 100.0%; Score 987; DB 8; Length 1567;
Best Local Similarity 100.0%; Pred. No. 7e-310;
Matches 987; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 ATGTCAGACACTTTGGCCACAACATGGAATCTCTCATCACACTGATGTTGACCCCTTCT 60
DB 537 ATGTCAGACACTTTGGCCACAACATGGAATCTCTCATCACACTGATGTTGACCCCTTCT 596
QY 61 GTCCTTCTCCCTCGGATCCAGTCTGGAACAAATTTTCATTTGCTGCTCTCACTCCCT 120
DB 597 GTCCTTCTCCCTCGGATCCAGTCTGGAACAAATTTTCATTTGCTGCTCTCACTCCCT 656
QY 121 GTGTGCTGCTTAGGCACAGCCACAATTTGGGCAATATAAATTAATCTGCTGTTGCTCC 180
DB 657 GTGTGCTGCTTAGGCACAGCCACAATTTGGGCAATATAAATTAATCTGCTGTTGCTCC 716
QY 181 ACTGAACAGCTTTCGACAGCCTGTGACCTTTTCTGTCATGCTCTCAACCATCGAC 240
DB 717 ACTGAACAGCTTTCGACAGCCTGTGACCTTTTCTGTCATGCTCTCAACCATCGAC 776
QY 241 TTGGCTGCTCTGTCTCCACAGTCTCCCAAGCTACTGGCTATCTTCTGCTGGAGCCGA 300
DB 777 TTGGCTGCTCTGTCTCCACAGTCTCCCAAGCTACTGGCTATCTTCTGCTGGAGCCGA 836
QY 301 CATATATCTGCTCTGCTGCTGCTGCGACATATGTTCTTCATTCATGCTCTTGCATGATG 360
DB 837 CATATATCTGCTCTGCTGCTGCTGCGACATATGTTCTTCATTCATGCTCTTGCATGATG 896
QY 361 GAGTCCACTGTCTACTGCGCATGGCTTTGATCGTACGTGGCCATCGCCACCCACTC 420
DB 897 GAGTCCACTGTCTACTGCGCATGGCTTTGATCGTACGTGGCCATCGCCACCCACTC 956
QY 421 CGCTATGCCACAATCTCTACTGACACCATCATTTGCCACATAGGGGTGGCAGCTGATG 480
DB 957 CGCTATGCCACAATCTCTACTGACACCATCATTTGCCACATAGGGGTGGCAGCTGATG 1016
QY 481 CGAGGCTCCCTGCTATGCTCCATGCTCCCTCTTTATTTGGGGTTTGAACCTTCTCCAA 540
DB 1017 CGAGGCTCCCTGCTATGCTCCATGCTCCCTCTTTATTTGGGGTTTGAACCTTCTCCAA 1076
QY 541 AGCCATGTATCTCTACACAGTACTGTGAGCAGCATGGCTGTGTGAGCTGGCTGTGGA 600
DB 1077 AGCCATGTATCTCTACACAGTACTGTGAGCAGCATGGCTGTGTGAGCTGGCTGTGGA 1136
QY 601 GACACAGGCTTAACCGTGTGTATGGGCTGACAGCTGCATGTTGCTCATTTGGGGTTGAC 660
DB 1137 GACACAGGCTTAACCGTGTGTATGGGCTGACAGCTGCATGTTGCTCATTTGGGGTTGAC 1196
QY 661 TTGTTTGTGATTTGCTCTCTATGCTTATGCTTATTTGCAAGCTGCTCTGCTCTCATCC 720
DB 1197 TTGTTTGTGATTTGCTCTCTATGCTTATTTGCAAGCTGCTCTGCTCTCATCC 1256
QY 721 CATGAAGCTCGGTCCAGGCGCTTAGGACCTGTGGTTCCCATGCTGTGTCTCATCTCATC 780

DB 1257 CATGAAGCTCGGTCCAGGCGCTTAGGACCTGTGGTTCCCATGCTGTGTCTCATCTCATC 1316
QY 781 TCTTATACACAGCCCT 840
DB 1317 TCTTATACACAGCCCT 1376
QY 841 CATATTACATCTCTTTTGGGCAATGTTATCTGCTTTTGGCCACCTGCTCTTAATCTCTGTG 900
DB 1377 CATATTACATCTCTTTTGGGCAATGTTATCTGCTTTTGGCCACCTGCTCTTAATCTCTGTG 1436
QY 901 GTATATGAGTTAAGACCAACACAGATCCGTAAGAGAGTTGTCAGGGTGTTCAGAGTGGG 960
DB 1437 GTATATGAGTTAAGACCAACACAGATCCGTAAGAGAGTTGTCAGGGTGTTCAGAGTGGG 1496
QY 961 CAGGAATGGCATCAAGGCATCTGAG 987
DB 1497 CAGGAATGGCATCAAGGCATCTGAG 1523
RESULT 2
AAS42339
ID AAS42339 standard; cDNA; 990 BP.
XX
AC AAS42339;
XX
DT 18-DEC-2001 (first entry)
XX
DE Human cDNA encoding olfactory receptor AOLFRL141.
XX
KW Human; olfactory receptor; G protein-coupled receptor; GPCR; odourant;
KW ss; food additive; cosmetic; fragrance; pharmaceutical additive.
XX
OS Homo sapiens.
XX
PN WO200168805-A2.
XX
PD 20-SEP-2001.
XX
PF 13-MAR-2001; 2001WO-US007771.
XX
PR 13-MAR-2000; 2000US-0188914P.
PR 24-MAR-2000; 2000US-0192033P.
PR 12-APR-2000; 2000US-0198474P.
PR 24-APR-2000; 2000US-0199335P.
PR 26-MAY-2000; 2000US-0207702P.
PR 23-JUN-2000; 2000US-0213849P.
PR 16-AUG-2000; 2000US-0226534P.
PR 07-SEP-2000; 2000US-0230732P.
PR 07-FEB-2001; 2001US-0266862P.
XX
PA (SENO-) SENOMYX INC.
XX
PI Zozulya S;
XX
DR WPI; 2001-570867/64.
DR P-PSDB; AAU24646.
XX
PT Nucleic acids encoding human olfactory G protein-coupled receptors, where
PT the compounds can be used in the food, pharmaceutical and cosmetic
PT industries to customize odors.
XX
PT
XX
PS Claim 1; Page 142; 319pp; English.
XX
CC The invention relates to nucleic acids encoding human olfactory
CC receptors, OR, (a G protein-coupled receptor, GPCR). The OR's
CC specifically recognise molecules, odourants, that elicit specific
CC olfactory sensation. The human olfactory receptors and polynucleotides
CC encoding them are useful for screening a library of chemical compounds
CC for compounds that are involved in olfactory sensation. Modulators of
CC their activity are useful for pharmacological and genetic modulation of
CC olfactory signalling pathways. Therefore, they can be used in the food,

CC pharmaceutical and cosmetic industries to customise odours and
CC fragrances. The present sequence encodes a human olfactory receptor of
CC the invention
XX
SQ Sequence 990 BP; 188 A; 285 C; 221 G; 296 T; 0 U; 0 Other;

Query Match 99.4%; Score 980.6; DB 5; Length 990;
Best Local Similarity 99.6%; Pred. No. 6.5e-308;
Matches 983; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

```
QY 1 ATGTCAGCACTCTTGGCCACAACTGGAATCTCTCATCACACTGATGTTGACCCCTTCT 60
Dd |||
QY 1 ATGTCAGCACTCTTGGCCACAACTGGAATCTCTCATCACACTGATGTTGACCCCTTCT 60
Dd |||
QY 61 GTCTTCTTCTCTCTGGGATCCAGGCTCTGGACATTTTCAATTTGGGCTCTCACTCCCT 120
Dd |||
QY 61 GTCTTCTTCTCTCTGGGATCCAGGCTCTGGACATTTTCAATTTGGGCTCTCACTCCCT 120
Dd |||
QY 121 GTGTGTGGCTTAGGCACAGCCCAATTTGTGGCAATATAACTATTCTGTTGTTGTTGCC 180
Dd |||
QY 121 GTGTGTGGCTTAGGCACAGCCCAATTTGTGGCAATATAACTATTCTGTTGTTGTTGCC 180
Dd |||
QY 181 ACTGAACCACTCTTGACAGAGCTGTGACCTTTTCTGTGATGCTCTCAACCATCGAC 240
Dd |||
QY 181 ACTGAACCACTCTTGACAGAGCTGTGACCTTTTCTGTGATGCTCTCAACCATCGAC 240
Dd |||
QY 241 TTGGCTGCTCTGTCTCCACAGTCTCCAGCTACTGGCTATCTTCTGTGTGGAGCCGGA 300
Dd |||
QY 241 TTGGCTGCTCTGTCTCCACAGTCTCCAGCTACTGGCTATCTTCTGTGTGGAGCCGGA 300
Dd |||
QY 301 CATATATCGCTCTGCTGCTGCGTGGCCACATATGTTCTTCAATCATGCTCTTGCATGATG 360
Dd |||
QY 301 CATATATCGCTCTGCTGCTGCGTGGCCACATATGTTCTTCAATCATGCTCTTGCATGATG 360
Dd |||
QY 361 GAGTCCACTGTCTACTGGCCATGCGCTTTGATCGCTACGTGGCCATCGCCACCACTC 420
Dd |||
QY 361 GAGTCCACTGTCTACTGGCCATGCGCTTTGATCGCTACGTGGCCATCGCCACCACTC 420
Dd |||
QY 421 CGCTATGCGCAATCTCTCACTGACACCATCATTTGCCACATAGGGGTGGAGCTGTAGTG 480
Dd |||
QY 421 CGCTATGCGCAATCTCTCACTGACACCATCATTTGCCACATAGGGGTGGAGCTGTAGTG 480
Dd |||
QY 481 CGAGGCTCCCTGCTCATGCTCCCATGCTCCCTTCTTATTGGGCGTTTCAACTTCTGCCAA 540
Dd |||
QY 481 CGAGGCTCCCTGCTCATGCTCCCATGCTCCCTTCTTATTGGGCGTTTCAACTTCTGCCAA 540
Dd |||
QY 541 AGCCATGTATCTTACACAGCTACTGTGAGCAGATGGCTGTGGTGAAGCTGGCTGTGGA 600
Dd |||
QY 541 AGCCATGTATCTTACACAGCTACTGTGAGCAGATGGCTGTGGTGAAGCTGGCTGTGGA 600
Dd |||
QY 601 GACACAGGCTTAACCGTGTATGGGCTGACAGCTGACACTGTTGTTGTTGGGTTGAC 660
Dd |||
QY 601 GACACAGGCTTAACCGTGTATGGGCTGACAGCTGACACTGTTGTTGTTGGGTTGAC 660
Dd |||
QY 661 TTGTTTTCATTTGTTCTCTCTATGCTCCCTAATTTGCAAGCTGCTTGGCTCTCATCC 720
Dd |||
QY 661 TTGTTTTCATTTGTTCTCTCTATGCTCCCTAATTTGCAAGCTGCTTGGCTCTCATCC 720
Dd |||
QY 721 CATGAAGCTGGTCAAGGCGCTTAGGACCTGTGGTTCCTATGCTGTGTCATCCTCATC 780
Dd |||
QY 721 CATGAAGCTGGTCAAGGCGCTTAGGACCTGTGGTTCCTATGCTGTGTCATCCTCATC 780
Dd |||
QY 781 TCTTATACACAGCCCTCTCTCTCTCTCTTTTACACACCGCTTTGGCCATCAGTTCAGTC 840
Dd |||
QY 781 TCTTATACACAGCCCTCTCTCTCTCTCTTTTACACACCGCTTTGGCCATCAGTTCAGTC 840
Dd |||
QY 841 CATATTACATTTCTTTTGGGCAATGTTTATCTGCTTTTGGCACTGCTTTAATCTCGTG 900
Dd |||
QY 841 CATATTACATTTCTTTTGGGCAATGTTTATCTGCTTTTGGCACTGCTTTAATCTCGTG 900
Dd |||
QY 901 GTATATGGAGTTAAGACCAACAGATCCGTAAGAGAGTTGTCAGGGTGTTCAAAGTGGG 960
Dd |||
QY 901 GTATATGGAGTTAAGACCAACAGATCCGTAAGAGAGTTGTCAGGGTGTTCAAAGTGGG 960
Dd |||
```

QY 961 CAGGGAATGGGCATCAAGGCATCTGAG 987
Dd |||
QY 961 CAGGGAATGGGCATCAAGGCATCTGAG 987
Dd |||

RESULT 3
ABZ43046
ID ABZ43046 standard; DNA; 990 BP.

XX ABZ43046;
AC ABZ43046;
DT 06-MAR-2003 (first entry)

XX Human GPCR polynucleotide SEQ ID NO 353.
DE Human; GPCR; G protein coupled receptor; signal transduction; olfactory;
XX Human; GPCR; gustatory; taste; fragrance; gene; ds.
KW drug development;
XX Homo sapiens.

OS WO200216548-A2.
PN 28-FEB-2002.
XX 30-JUL-2001; 2001WO-IB001446.
PF 04-AUG-2000; 2000JP-00237818.
PR 13-FEB-2001; 2001JP-00034434.

XX (NISC-) JAPAN SCI & TECHNOLOGY CORP.
PA Haga T, Takeda S, Mitaku S;
XX WPI; 2002-304118/34.
DR P-PSDB; ABP95772.

XX Database global search for G protein-coupled receptors, proteins and
PT encoded genes for studying in vivo signal transduction mechanism and
PT identifying targets for drug development.

XX Claim 9; SEQ ID NO 353; 97pp + Sequence Listing; Japanese.
XX The invention relates to a method for screening G protein-coupled
CC receptor (GPCR) genes (ABZ42870-ABZ43216) and/or GPCR proteins (ABP95596-
CC ABP95942) by extracting open-reading frames containing 6-8 transmembrane
CC domains with 250-1000 amino acid residues to give a gene homologous with
CC a known GPCR gene. The receptor proteins and encoded genes are useful for
CC studying in vivo signal transduction mechanism and identifying targets
CC for drug development e.g. based on olfactory and gustatory receptors in
CC form of agonists and antagonists by screening intrinsic and extrinsic
CC ligands as bitter taste inhibitors, taste enhancers and fragrance
CC improvers. Note: The sequence data for this patent did not form part of
CC the printed specification, but was obtained in electronic format directly
CC from WIPO at ftp.wipo.int/pub/published_pct_sequences

XX SQ Sequence 990 BP; 188 A; 285 C; 221 G; 296 T; 0 U; 0 Other;
Query Match 99.4%; Score 980.6; DB 6; Length 990;
Best Local Similarity 99.6%; Pred. No. 6.5e-308;
Matches 983; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 ATGTCAGCACTCTTGGCCACAACTGGAATCTCTCATCACACTGATGTTGACCCCTTCT 60
Dd |||
QY 1 ATGTCAGCACTCTTGGCCACAACTGGAATCTCTCATCACACTGATGTTGACCCCTTCT 60
Dd |||
QY 61 GTCTTCTTCTCTCTGGGATCCAGGCTCTGGACATTTTCAATTTGGGCTCTCACTCCCT 120
Dd |||
QY 61 GTCTTCTTCTCTCTGGGATCCAGGCTCTGGACATTTTCAATTTGGGCTCTCACTCCCT 120
Dd |||
QY 121 GTGTGTGGCTTAGGCACAGCCCAATTTGTGGCAATATAACTATTCTGTTGTTGTTGCC 180
Dd |||
QY 121 GTGTGTGGCTTAGGCACAGCCCAATTTGTGGCAATATAACTATTCTGTTGTTGTTGCC 180
Dd |||

Db 122 GTGTGGCTTAGGCACAGCCACAATTTGTGGGCAATATAACTATTCTGGTTGTGTGCC 181
Qy 181 ACTGAACCAAGCTTTGACAAAGCCGTGTACCTTTTCTGTGCATGCTCTCAACCATCGAC 240
Db 182 ACTGAACCAAGCTTTGACAAAGCCGTGTACCTTTTCTGTGCATGCTCTCAACCATCGAC 241
Qy 241 TTGGCTGCCCTCTGTCTCCACAGTTCCCAAGTACTGGCTATCTCTGGTGTGGAGCCGGA 300
Db 242 TTGGCTGCCCTCTGTCTCCACAGTTCCCAAGTACTGGCTATCTCTGGTGTGGAGCCGGA 301
Qy 301 CATATATCTGCCCTGCTGCTGCGCATATGTTCTTCACTCATGCTCTCTGCATGATG 360
Db 302 CATATATCTGCCCTGCTGCTGCGCATATGTTCTTCACTCATGCTCTCTGCATGATG 361
Qy 361 GAGTCCACTGTGCTAGTGGCCATGGCCCTTTGATCGCTACGTGGCCATTCGCCACCCACTC 420
Db 362 GAGTCCACTGTGCTAGTGGCCATGGCCCTTTGATCGCTACGTGGCCATTCGCCACCCACTC 421
Qy 421 CGTATATGCCAATCTCACTGACACCATCATTTGCCACATAGGGTGGAGCTGTAGTG 480
Db 422 CGTATATGCCAATCTCACTGACACCATCATTTGCCACATAGGGTGGAGCTGTAGTG 481
Qy 481 CGAGGCTCCCTGCTCATGCTCCCATGCTCCCTTTTATTGGGCGTTTGAACCTTCTGCCAA 540
Db 482 CGAGGCTCCCTGCTCATGCTCCCATGCTCCCTTTTATTGGGCGTTTGAACCTTCTGCCAA 541
Qy 541 AGCCATGTGATCTTACACACATGCTGTGAGCACATGGCTGTGGTGAAGCTGCCCTGTGGA 600
Db 542 AGCCATGTGATCTTACACACATGCTGTGAGCACATGGCTGTGGTGAAGCTGCCCTGTGGA 601
Qy 601 GACACCAAGCCCTAACCGGTGTATGGGTGTGACAGCTGCTGTGGTCAATGGGGTTGAC 660
Db 602 GACACCAAGCCCTAACCGGTGTATGGGTGTGACAGCTGCTGTGGTCAATGGGGTTGAC 661
Qy 661 TTGTTTTCATTTGCTCTCTATGCTCCCTTAATGACAAAGCTGTCTTCCGCTCTCATCC 720
Db 662 TTGTTTTCATTTGCTCTCTATGCTCCCTTAATGACAAAGCTGTCTTCCGCTCTCATCC 721
Qy 721 CATGAAGCTCGGTCAAGGCCCTTAGGGACCTGTGGTTCCTCATGCTGTGTCACTCATC 780
Db 722 CATGAAGCTCGGTCAAGGCCCTTAGGGACCTGTGGTTCCTCATGCTGTGTCACTCATC 781
Qy 781 TCTTATACACAGCCCTCTCTCTCTTTTACACACCGCTTTGGCCATCAGCTTCCAGTC 840
Db 782 TCTTATACACAGCCCTCTCTCTCTTTTACACACCGCTTTGGCCATCAGCTTCCAGTC 841
Qy 841 CATATTTCATTTTGGCCAAATGTTTATCTGCTTTTGGCCACCTGCTTAACTCTGTG 900
Db 842 CATATTTCATTTTGGCCAAATGTTTATCTGCTTTTGGCCACCTGCTTAACTCTGTG 901
Qy 901 GTATATGGAGTTAAGACCAACAGATCCGTAAGAGAGTTGTCAAGGTGTTTCAAGTGGG 960
Db 902 GTATATGGAGTTAAGACCAACAGATCCGTAAGAGAGTTGTCAAGGTGTTTCAAGTGGG 961
Qy 961 CAGGGAATGGGATCAAGGCATCTGAG 987
Db 962 CAGGGAATGGGATCAAGGCATCTGAG 988

RESULT 7
ADC86344
ID ADC86344 standard; DNA; 1390 BP.
XX AC
AC ADC86344;
XX AC
DT 01-JAN-2004 (first entry)
XX DE
DE Human GPCR gene SEQ ID NO:797.
XX KW
KW ds; gene; human; GPCR;
KW guanosine triphosphate-binding protein coupled receptor; gene therapy.
XX OS
OS Homo sapiens.

XX PN EP1270724-A2.
XX PD 02-JAN-2003.
XX PF 18-JUN-2002; 2002EP-00013517.
XX PR 18-JUN-2001; 2001JP-00246789.
XX PA (NAAD-) NAT INST ADVANCED IND SCI & TECHNOLOGY.
XX PA (ADSC-) CENT ADVANCED SCI & TECHNOLOGY INCUBATIO.
XX PI Suwa M, Asai K, Akiyama Y, Aburatani H;
XX MPI; 2003-315783/31.
XX DR P-PSDB; ADC86345.
XX PT New polynucleotide, useful for preparing a composition for treating a
PT patient in need of increased or suppressed activity or expression of the
PT guanosine triphosphate-binding protein coupled receptor.
XX FS Claim 1; SEQ ID NO 797; 28pp; English.
XX CC The invention relates to a novel polynucleotide encoding a guanosine
CC triphosphate-binding protein coupled receptor (GPCR). A polynucleotide of
CC the invention may have a use in gene therapy. The polynucleotide and
CC polypeptide are useful for preparing a composition for treating a patient
CC in need of increased or suppressed activity or expression of the
CC guanosine triphosphate-binding protein coupled receptor. The
CC polynucleotide sequences shown in ADC85548-ADC87616 encode GPCR's of the
XX invention.
XX SQ Sequence 1390 BP; 317 A; 353 C; 293 G; 427 T; 0 U; 0 Other;
Query Match 99.4%; Score 980.6; DB 9; Length 1390;
Best Local Similarity 99.6%; Pred. No. 7.9e-308;
Matches 983; Conservative 0; Mismatches 4; Indels 0; Gaps 0;
Qy 1 ATGTCCAGCACTCTTGGCCACACATGGAATCTCCTCATCACACTGATGTTGACCTTCT 60
Db 201 ATGTCCAGCACTCTTGGCCACACATGGAATCTCCTCATCACACTGATGTTGACCTTCT 260
Qy 61 GTCTTCTTCTCTGGGCATCCAGGTCTGGAAACAATTTCAATTTTGGTCTCACTCCCT 120
Db 261 GTCTTCTTCTCTGGGCATCCAGGTCTGGAAACAATTTCAATTTTGGTCTCACTCCCT 320
Qy 121 GTGTGTGCTTAGGCACAGCCACAATTTGGGCAATAAATACTTCTGGTGTGTGTGCC 180
Db 321 GTGTGTGCTTAGGCACAGCCACAATTTGGGCAATAAATACTTCTGGTGTGTGTGCC 380
Qy 181 ACTGAACCAAGCTTTGACAAAGCCGTGTACCTTTTCTGTGCATGCTCTCAACCATCGAC 240
Db 381 ACTGAACCAAGCTTTGACAAAGCCGTGTACCTTTTCTGTGCATGCTCTCAACCATCGAC 440
Qy 241 TTGGTGTGCTCTGTCTCCACAGTTCCCAAGTACTGGCTATCTTCTGTGTGTGGAGCCGA 300
Db 441 TTGGTGTGCTCTGTCTCCACAGTTCCCAAGTACTGGCTATCTTCTGTGTGTGGAGCCGA 500
Qy 301 CATATATCTGCTCTGCCCTGCTGCGCATATGTTCTTCACTCATGCTCTCTGCATGATG 360
Db 501 CATATATCTGCTCTGCCCTGCTGCGCATATGTTCTTCACTCATGCTCTCTGCATGATG 560
Qy 361 GAGTCCACTGTGCTAGTGGCCATGGCCCTTTGATCGCTACGTGGCCATTCGCCACCCACTC 420
Db 561 GAGTCCACTGTGCTAGTGGCCATGGCCCTTTGATCGCTACGTGGCCATTCGCCACCCACTC 620
Qy 421 CGCTATGCCAATCTCTCACTGACACCATCATTTGCCACATAGGGTGGAGCTGTAGTG 480
Db 621 CGCTATGCCAATCTCTCACTGACACCATCATTTGCCACATAGGGTGGAGCTGTAGTG 680
Qy 481 CGAGGCTCCCTGCTCATGCTCCCATGCTCCCTTCTTTATTGGGCGTTTGAACCTTCTGCCAA 540
Db 681 CGAGGCTCCCTGCTCATGCTCCCATGCTCCCTTCTTTATTGGGCGTTTGAACCTTCTGCCAA 740

RESULT 11
ABZ777974
ID ABZ77974 standard; DNA; 948 BP.
XX AC
XX ABZ77974;
DT 17-APR-2003 (first entry)
XX
DE Human G protein coupled receptor coding sequence SEQ ID 205.
XX
XX Human; gene; anorectic; antiasthmatic; antidiabetic; hypotensive;
KW antiparkinsonian; nootropic; neuroprotective; tranquiliser;
KW antirheumatic; antiinflammatory; osteopathic; cardiant; neuroleptic;
KW antiarthritic; gene therapy; olfactory G protein-coupled receptor; GPCR;
KW infection; obesity; diabetes; hypertension; malnutrition;
KW Parkinson's disease; Alzheimer's disease; Korsakoff's psychosis; anxiety;
KW rheumatoid arthritis; chronic obstructive pulmonary disease;
KW osteoporosis; asthma; myocardial infarction; schizophrenia;
KW osteoarthritis; ds.
XX
OS Homo sapiens.
XX
XX WO2003000735-A2.
XX
XX 03-JAN-2003.
XX
XX 24-JUN-2002; 2002WO-IB002481.
XX
XX 26-JUN-2001; 2001US-0301095P.
XX
XX 06-NOV-2001; 2001US-0332758P.
XX
XX (DECO-) DECODE GENETICS EHF.
XX
XX Martinez RAM, Sigurdsson GT;
XX
XX WPI; 2003-175284/17.
XX
XX P-PSDB; ABR01673.
XX
XX New olfactory G protein-coupled receptor gene nucleic acid and
PT polypeptide, useful for diagnosing or treating a disease or condition
PT associated with GPCR, e.g. obesity, diabetes, hypertension, malnutrition
PT or Alzheimer's disease.
XX
XX Claim 1; Page 103; 383pp; English.
XX
XX The present invention relates to novel human olfactory G protein-coupled
CC receptors (GPCR) and their coding sequences (ABZ77872-ABZ77986 and
CC ABR01571-ABR01685). The GPCRs and coding sequences are useful for
CC diagnosing or treating a disease or condition associated with GPCR, e.g.
CC infections, obesity, diabetes, hypertension, malnutrition, Parkinson's
CC disease, Alzheimer's disease, Korsakoff's psychosis, anxiety, rheumatoid
CC arthritis, chronic obstructive pulmonary disease, osteoporosis, asthma,
CC myocardial infarction, schizophrenia, or osteoarthritis
XX
SQ Sequence 948 BP; 174 A; 266 C; 202 G; 306 T; 0 U; 0 Other;

Query Match 80.1%; Score 791; DB 7; Length 948;
Best Local Similarity 94.3%; Pred. No. 3.2e-246;
Matches 821; Conservative 0; Mismatches 50; Indels 0; Gaps 0;

QY 58 TCTGCTTCTTCTCTCTGGGATCCAGGCTGGGAACATTTTCATTTGGGCTCACTC 117
DB 28 TCTGAGTTATCTCTGATGGATCCCTGGCAATTCACATTTCTGTTTTTTT 87
QY 118 CTTGCTGTGGCTTAGGCACAGCCAAATTTGGGCAATAAATATTTCTGTTCTT 177
DB 86 TTGTTGTTGTTGTTGTTTCTTCCAAATTTGGGCAATAAATATTTCTGTTGTT 147
QY 178 GCCACTGAACCACTTTCACAAAGCCTGTGTACCTTTTCTGTGTGATGCTTCAACCATC 237
DB 148 GCCACTGAACCACTTTCACAAAGCCTGTGTACCTTTTCTGTGTGATGCTTCAACCATC 207

QY 238 GACTTGGCTGCTCTGCTCTCCACAGTTCCCAAGCTACTGGCTATCTTCTGTTGTGGAGCC 297
DB 208 GACTTGGCTGCTCTGCTCTCCACAGTTCCCAAGCTACTGGCTATCTTCTGTTGTGGAGCC 267
QY 298 GGACATATATCTGCTCTGCTGCTGGCAGATATGTTCTTCAATTCATGCTCTTCTGCAATG 357
DB 268 GGACATATATCTGCTCTGCTGCTGGCAGATATGTTCTTCAATTCATGCTCTTCTGCAATG 327
QY 358 ATGGAGTCCACTGTGCTACTGGCCATGGCCCTTTGATCGCTACGTGGCCATCTGCCACCCA 417
DB 328 ATGGAGTCCACTGTGCTACTGGCCATGGCCCTTTGATCGCTACGTGGCCATCTGCCACCCA 387
QY 418 CTCGCTATGCGACAAATCTCTCACTGACACCAATCATTTGCCACATAGGGGTGGCACTGTA 477
DB 388 CTCGCTATGCGACAAATCTCTCACTGACACCAATCATTTGCCACATAGGGGTGGCACTGTA 447
QY 478 GTGGAGGCTCCCTGCTCATGCTCCCATGTCCTTCTTTTATTTGGGCGTTTGAATTTCTGC 537
DB 448 GTGGAGGCTCCCTGCTCATGCTCCCATGTCCTTCTTTATTTGGGCGTTTGAATTTCTGC 507
QY 538 CAAAGCCATGTATCTCTACACACAGTACTGTGAGCACATGGCTGTGGTGAAGCTGGCCTGT 597
DB 508 CAAAGCCATGTATCTCTACACACAGTACTGTGAGCACATGGCTGTGGTGAAGCTGGCCTGT 567
QY 598 GGAGACACAGGCTTAACCGTGTGTATGGGCTGACAGCTGCACCTGTTGGTCAATTTGGGTT 657
DB 568 GGAGACACAGGCTTAACCGTGTGTATGGGCTGACAGCTGCACCTGTTGGTCAATTTGGGTT 627
QY 658 GACTTGTGTTTGCATTTGCTCTCTCTATGCCCTTAATTTGACAAAGCTGTCTTGGCCTCTCA 717
DB 628 GACTTGTGTTTGCATTTGCTCTCTCTATGCCCTTAATTTGACAAAGCTGTCTTGGCCTCTCA 687
QY 718 TCCCATGAAGCTCGGTCACAGGCGCTTAGGACCTGTGTGTTCCCATGTCTGTGTATCTTC 777
DB 688 TCCCATGAAGCTCGGTCACAGGCGCTTAGGACCTGTGTGTTCCCATGTCTGTGTATCTTC 747
QY 778 ATCTCTTATACACAGGCGCTCTCTCTCTTTTATACACAGGCTTTGGGCAATCACCTTCCA 837
DB 748 ATCTCTTATACACAGGCGCTCTCTCTCTTTTATACACAGGCTTTGGGCAATCACCTTCCA 807
QY 838 GTCCATATTTCAATTTCTTTTGGGCAATGTTTATCTGCTTTTGGCACCCTGCTTAACTCT 897
DB 808 GTCCATATTTCAATTTCTTTTGGGCAATGTTTATCTGCTTTTGGCACCCTGCTTAACTCT 867
QY 898 GTGGTATATGGAGTTAAGACCAACAGATCC 928
DB 868 GTGGTATATGGAGTTAAGACCAACAGATCC 898
RESULT 12
AAD19143
ID AAD19143 standard; cDNA; 980 BP.
XX AC
XX AAD19143;
DT 18-DEC-2001 (first entry)
XX
XX Human G-protein coupled receptor 12 (GPCR12) cDNA.
XX
XX Human; G-protein coupled receptor 12; GPCR12; cardiomyopathy; vaccine;
KW atherosclerosis; diabetes; cardiant; cyclostatic; cancer; obesity; pain;
KW diabetes mellitus; anorexia; cachexia; cardiomyopathy; atherosclerosis;
KW neurodegenerative disorder; Alzheimer's disease; Parkinson's disease;
KW immune disorder; haematopoietic disorders; retinal disorder; HIV;
KW human immunodeficiency virus; adenocarcinoma; bulimia; asthma; ulcer;
KW angina pectoris; hypotension; hypertension; Crohn's disease; anxiety;
KW multiple sclerosis; schizophrenia; dementia; mental retardation;
KW gene therapy; osteoporosis; urinary retention; ss.
XX
OS Homo sapiens.
XX
XX Key Location/Qualifiers
XX PH 15. .959
XX CDS

This Page Blank (u.p.10)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 27, 2004, 16:11:37 ; Search time 93 Seconds
(without alignments)
5889.640 Million cell updates/sec

Title: US-10-081-775-1_COPY_537_1523

Perfect score: 987
Sequence: 1 atgtccagcactcttgcca.....tggscatcaaggaatctgag 987

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 682709 seqs, 277475446 residues

Total number of hits satisfying chosen parameters: 1365418

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued Patents NA.*

- 1: /cgn2_6/ptodata/2/ina/5A COMB.seq.*
- 2: /cgn2_6/ptodata/2/ina/5B COMB.seq.*
- 3: /cgn2_6/ptodata/2/ina/6A COMB.seq.*
- 4: /cgn2_6/ptodata/2/ina/6B COMB.seq.*
- 5: /cgn2_6/ptodata/2/ina/PTUS COMB.seq.*
- 6: /cgn2_6/ptodata/2/ina/backfiles1.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	334	33.8	963	4	US-09-439-313-526 Sequence 526, App
2	334	33.8	963	4	US-09-636-215-526 Sequence 526, App
3	334	33.8	963	4	US-09-685-166A-526 Sequence 526, App
4	330.8	33.5	1474	1	US-08-465-980-1 Sequence 1, Appli
5	330.8	33.5	1474	2	US-09-053-303-1 Sequence 1, Appli
6	330.8	33.5	1474	4	US-09-339-115-1 Sequence 1, Appli
7	330.8	33.5	1474	5	PCT-US95-07093-1 Sequence 1, Appli
8	326.2	33.0	1539	4	US-09-668-680-13 Sequence 13, Appli
9	142.4	14.4	1854	4	US-09-016-434-1312 Sequence 1312, Ap
10	137.2	13.9	966	3	US-08-748-506-5 Sequence 5, Appli
11	133.2	13.5	966	3	US-08-748-506-7 Sequence 7, Appli
12	131.8	13.4	966	3	US-08-748-506-6 Sequence 6, Appli
13	131	13.3	1713	2	US-08-467-948A-1 Sequence 1, Appli
14	131	13.3	1713	3	US-08-467-947A-1 Sequence 1, Appli
15	126.8	12.8	966	3	US-08-748-506-8 Sequence 8, Appli
16	118.4	12.0	1990	4	US-09-016-434-1056 Sequence 1056, Ap
17	117.6	11.9	675	4	US-09-465-901-43 Sequence 43, Appli
18	117	11.9	1282	4	US-09-016-434-1413 Sequence 1413, Ap
19	116.8	11.8	3459	4	US-09-016-434-1363 Sequence 1363, Ap
20	115.4	11.7	1035	4	US-09-546-986A-1 Sequence 1, Appli
21	115.4	11.7	1035	4	US-09-524-730-1 Sequence 1, Appli
22	109	11.0	1062	4	US-09-668-680-10 Sequence 10, Appli
23	107	10.8	678	4	US-09-465-901-45 Sequence 45, Appli
24	106.4	10.8	1080	4	US-09-668-680-9 Sequence 9, Appli
25	105.8	10.7	1351	4	US-09-546-986A-5 Sequence 5, Appli
26	105.8	10.7	1351	4	US-09-524-730-5 Sequence 5, Appli
27	102.2	10.4	1438	4	US-09-016-434-1313 Sequence 1313, Ap

28	101.2	10.3	666	4	US-09-465-901-33 Sequence 33, Appli
29	99.6	10.1	542	4	US-09-016-434-627 Sequence 627, App
30	99.4	10.1	1065	4	US-03-546-986A-7 Sequence 7, Appli
31	99.4	10.1	1065	4	US-09-524-730-7 Sequence 7, Appli
32	92.2	9.3	234	4	US-09-016-434-1047 Sequence 1047, Ap
33	91.4	9.3	675	4	US-09-465-901-17 Sequence 17, Appli
34	90.6	9.2	669	4	US-09-465-901-25 Sequence 25, Appli
35	89	9.0	669	4	US-09-465-901-11 Sequence 11, Appli
36	88.8	9.0	648	4	US-09-016-434-1374 Sequence 1374, Ap
37	87.2	8.8	648	4	US-09-016-434-1371 Sequence 1371, Ap
38	85.6	8.7	900	3	US-09-085-371-5 Sequence 5, Appli
39	84.2	8.5	645	4	US-09-016-434-1372 Sequence 1372, Ap
40	82.2	8.3	648	4	US-09-016-434-1375 Sequence 1375, Ap
41	81.6	8.3	1290	2	US-08-827-291A-1 Sequence 1, Appli
42	79.2	8.0	675	4	US-09-016-434-1373 Sequence 1373, Ap
43	78.6	8.0	648	4	US-09-016-434-1376 Sequence 1376, Ap
44	78.4	7.9	669	4	US-09-465-901-13 Sequence 13, Appli
45	75.4	7.6	952	4	US-09-016-434-1115 Sequence 1115, Ap

ALIGNMENTS

RESULT 1

US-09-439-313-526
; Sequence 526, Application US/09439313

; Patent No. 6329505

; GENERAL INFORMATION:

; APPLICANT: Xu, Jiangchun

; APPLICANT: Dillon, Davin C.

; APPLICANT: Mitcham, Jennifer L.

; APPLICANT: Harlocker, Susan Louise

; APPLICANT: Jiang Yuqi

; APPLICANT: Reed, Steven G.

; APPLICANT: Kalos, Michael

; APPLICANT: Fanger, Gary

; APPLICANT: Retter, Mark

; APPLICANT: Solk, John

; APPLICANT: Day, Craig

; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THERAPY AND

; TITLE OF INVENTION: DIAGNOSIS OF PROSTATE CANCER

; FILE REFERENCE: 210121.427C9

; CURRENT APPLICATION NUMBER: US/09/439,313

; NUMBER OF SEQ ID NOS: 575

; SOFTWARE: FastSeq for Windows Version 3.0

; SEQ ID NO 526

; LENGTH: 963

; TYPE: DNA

; ORGANISM: Homo sapiens

US-09-439-313-526

Query Match 33.8%; Score 334; DB 4; Length 963;

Best Local Similarity 60.2%; Pred. No. 1.6e-99;

Matches 553; Conservative 0; Mismatches 365; Indels 0; Gaps 0;

QY	63	CTTCTTCCTCGGCACTCCAGGCTCGAAACAAATTTGCTGCTCTACTCCCTGT	122
DB	30	CTTTGCTTATTGTTATCCCAAGATTAGAGAAAGCCCAATTTCTGGTTGGCTTCC	89
QY	123	GTGTGGCTTAGGCACAGCCACAAATTTGGGCAATATAACTATTCTGTTGTTGGCCAC	182
DB	90	CTTTCCATGATGATGAGGCAATGTTTGGAACTGCATCGTGGTCTTCATCGTAAGGAC	149
QY	183	TGAACCAAGTCTTGCAAGCCCTGTGTACTTTTCTGTGATGCTCTCAACCATGACTT	242
DB	150	GGACGACGCTTGACGCTCCGATGACCTTTCTCTCATGCTTGACGCAATTCACCT	209
QY	243	GGCTGCTCTGTCTCCAGAGTTCACAGTCTCCAGTACTGGTATCTTCTGTTGGAGCCGACA	302
DB	210	GGCTTTATCCATCCACCATCCACATCCCTAAAGATCCTTTCGCTTTTCTGGTTGATCCGAGA	269
QY	303	TATATCTGCTCTGCTGCTGCTGCACATATGTTCTTCATTCATGCTTCTGCTGATGGA	362

Db 270 GATTAGCTTTGAGGCTGCTGCTTACCCAGATGTTCTTTATTCATGCCCTCTCAGCCATTGA 329
Qy 363 GTCCACTGTGCTACTGGCCATGGCTTTGATGCTAGCTGGCCATCTGCCACCCACTCG 422
Db 330 ATCCACCACTCTGCTGGCCATGGCTTTGACCGTTATGTGGCCATCTGCCACCCACTCG 389
Qy 423 CTATGCCAATCTCTCACTGACACCATCATTCGCCACATAGGGGTGGCAGCTGTATGGC 482
Db 390 CCAATGCTGAGTGTCTCAACATACAGTAACAGCCCAAGATTGGCATCGTGGCTGTGTCG 449
Qy 483 AGGTCCTCTGCTCATGCTCCCATGTCCTTTTATTGGGCGTTTGAACCTTCTGCCAAG 542
Db 450 CGGATCCCTCTTTTTCCTCACTGCTCTGCTGATCAAGCGCTGGCCCTTCTGCCACTC 509
Qy 543 CCATGATGCTTACACAGCTACTGTGAGCAGATGGCTGTGGTGAAGCTGGCTGTGGAGA 602
Db 510 CAATGCTCTCTCGCATCTCTATGTGTCCACAGGATGTAATGAAGTTGGCTTATGCAGA 569
Qy 603 CACAGGCGCTAACCGTGTGTATGGGCTGACAGCTGCACCTGTTGGTCAATGGGTTGACTT 662
Db 570 CACTTTGCCAATGTGGTATATGTTCTTACTGCTTCTGCTGCTGCTGCTGCTGCTGCT 629
Qy 663 GTTTTGCATGTGCTCTATGCTCCCTAATTCGACAGCTGTCTTGGCTCTCATPCCA 722
Db 630 AATGTTCTATCTCTTGTCTTATTTCTGATAATACGACGGTCTGCAACTGCTCTCCAA 689
Qy 723 TGAAGCTGGTCCAGGCGCTAGGACCTGTGGTCCCATGCTGTGTGCTATCTCATCTC 782
Db 690 GTACAGCGGGCGCAAGGCTTTGGAACTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 749
Qy 783 TTATACACAGCCCTCTCTCTCTTTTACACACCGCTTTGGCCATCAGTTCCAGTCCA 842
Db 750 CTATGTGCCACTTATGGCCCTCAGTTGTATACACCGCTTTGGAAAGCCCTTCATCCCAT 809
Qy 843 TATTACATCTTTTGGCCAAATGTTATCTGCTTTTGGCACCTGCTCTTAATPCTGTGGT 902
Db 810 TGTGCTGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 869
Qy 903 ATATGGAGTTAAGCAACAGATCCGTAAGAGTTGTCAAGGTGTGTCAAGGTGGGCA 962
Db 870 CTATGTGCCAAACCAACAGATCAGAACCGGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 929
Qy 963 GGAATGGGCTATCAAGGC 980
Db 930 TGACAAGGACTTGAGGC 947

RESULT 2

US-09-636-215-526

; Sequence 526, Application US/09636215

; Patent No. 6620922

; GENERAL INFORMATION:

; APPLICANT: Xu, Jiangchun

; APPLICANT: Dillon, Devin C.

; APPLICANT: Mitcham, Jennifer L.

; APPLICANT: Harlocker, Susan L.

; APPLICANT: Jiang, Yuqi

; APPLICANT: Henderson, Robert A.

; APPLICANT: Kalos, Michael D.

; APPLICANT: Fanger, Gary R.

; APPLICANT: Retter, Marc W.

; APPLICANT: Stolk, John A.

; APPLICANT: Day, Craig H.

; APPLICANT: Vedvick, Thomas S.

; APPLICANT: Carter, Derrick

; APPLICANT: Li, Samuel

; APPLICANT: Wang, Aijun

; APPLICANT: Skeiky, Yasir A.W.

; APPLICANT: Hepler, William

; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND

; TITLE OF INVENTION: DIAGNOSIS OF PROSTATE CANCER

; FILE REFERENCE: 210121.42717C17

; CURRENT APPLICATION NUMBER: US/09/636,215

; CURRENT FILING DATE: 2000-08-10

; NUMBER OF SEQ ID NOS: 852

; SOFTWARE: FastSeq for Windows Version 3.0

; SEQ ID NO 526

; LENGTH: 963

; TYPE: DNA

; ORGANISM: Homo sapiens

US-09-636-215-526

Query Match

33.8%; Score 334; DB 4; Length 963;

Best Local Similarity 60.2%; Pred. No. 1.6e-99;

Matches 553; Conservative 0; Mismatches 365; Indels 0; Gaps 0;

Qy 63 CTTTCCTCTCTGGGCATCCAGGCTCTGGAAACAATTTATTTGTGGCTCTCATCTCCTCTG 122

Db 30 CTTTGTGCTTATTTGTTATCCAGGATTAGAGAAAGCCCATTTCTGGTGTGGCTTCCCTCT 89

Qy 123 GTGTGGCTTAGGCACAGCCAGCAATTTGTGGGCAATATACTATTCTGGTGTGTGGCCAC 182

Db 90 CTTTCCCATGTATGTAGTGGCAATTTTGGAAACTGCATCGTGGTCTTTCATCGTAAGGAC 149

Qy 183 TGAACCAAGTCTTGACAAAGCCTGTGTACCTTTTCTGTGCATGCTCTCAACCATCGACTT 242

Db 150 GGAACGCGAGCTGCAGGCTCCGATGCTCTTTCTGTGCATGCTCTGAGCCATTTGACCT 209

Qy 243 GGTCTCTCTGTCTCCACAGTTCACAGTCTACTGGCTATCTTCTGTGTGGAGCCGGACA 302

Db 210 GGCCTTATCCACATCCACCATGCTTAAGATCTTCTGCTGCTGCTGAGCCATTTGACCT 269

Qy 303 TATATCTGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 362

Db 270 GATTAGCTTTGAGGCTCTGCTTACCCAGATGTTCTTTATTCATGCTCTCTCAGCCATTGA 329

Qy 363 GTCCACTGTGCTACTGGCCATGGCTTTGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 422

Db 330 ATCCACCATCTGCTGGCCATGGCTTTGACCGTTATGTGGCCATCTGCCACCCACTGG 389

Qy 423 CTATGCCAATCTCTCACTGACACCATCATTTGCCACATAGGGGTGGCAGCTGTAGTGG 482

Db 390 CCATGCTGCTGCTCAACAATACAGTAAAGCCAGATGGCATGCTGCTGCTGCTGCTGCTG 449

Qy 483 AGGCTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 542

Db 450 CGGATCCCTCTTTTTCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 509

Qy 543 CCATGCTGCTTACACAGCTACTGTGACCATGGCTGTGCTGCTGCTGCTGCTGCTGCTGCTG 602

Db 510 CAATGCTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 569

Qy 603 CACGAGGCTTAACCGTGTGTATGGGCTGACAGCTGCACTGTTGTGCTGCTGCTGCTGCTG 662

Db 570 CACTTGGCCAAATGTGGTATATGCTTACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 629

Qy 663 GTTTTGCATGCTCTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 722

Db 630 AATGTTCACTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 689

Qy 723 TGAAGCTCGGTCCAAAGGCTCTAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 782

Db 690 GTACAGCGGGCCAAAGGCTTTGGAACTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTG 749

Qy 783 TTATACAGGCT 842

Db 750 CTATGTGCCAATTTATGGCCCTCTCAGTTGTACACCGCTTTGGAAACAGCTTCTCATCC 809

Qy 843 TATTACATCTTTTGGCCAAATGTTATCTGCTTTTGGCCACTGCTCTCTTAATCTCTGCTG 902

Db 810 TGTGCTGTTGTCTATGGGTGACATCTACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 869

Qy 903 ATATGGAGTTAAGCAACAGATCCGTAAGAGTTGTGAGGTTGTTTCAAGTGGGCA 962

Db 870 CTATGTGCCAAACCAACAGATCAGAACCGGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCT 929

QY 963 GGGAAATGGGCATCAAGGC 980
Db 930 TGACAAGGACTTGCAGGC 947

RESULT 3
US-09-685-166A-526
; Sequence 526, Application US/09685166A
; Patent No. 6630305
; GENERAL INFORMATION:
; APPLICANT: Xu, Jiangchun
; APPLICANT: Dillon, Davin C.
; APPLICANT: Mitcham, Jennifer L.
; APPLICANT: Harlocker, Susan L.
; APPLICANT: Jiang, Yuqi
; APPLICANT: Henderson, Robert A.
; APPLICANT: Kalos, Michael D.
; APPLICANT: Fanger, Gary R.
; APPLICANT: Retter, Marc W.
; APPLICANT: Stolk, John A.
; APPLICANT: Day, Craig H.
; APPLICANT: Vedwick, Thomas S.
; APPLICANT: Carter, Darrick
; APPLICANT: Li, Samuel
; APPLICANT: Wang, Aijun
; APPLICANT: Skeiky, Yasir A.W.
; APPLICANT: Hepler, William
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
; TITLE OF INVENTION: DIAGNOSIS OF PROSTATE CANCER
; FILE REFERENCE: 210121.427C21
; CURRENT APPLICATION NUMBER: US/09/685,166A
; CURRENT FILING DATE: 2000-10-10
; NUMBER OF SEQ ID NOS: 898
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 526
; LENGTH: 963
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-685-166A-526

Query Match 33.8%; Score 334; DB 4; Length 963;
Best Local Similarity 60.2%; Pred. No. 1.6e-99;
Matches 553; Conservative 0; Mismatches 365; Indels 0; Gaps 0;

QY 63 CTTCTTCTCTCTGGGATCCAGGCTGGAAACAATTTCAATTTGTGGCTCTCACTCCCTGT 122
Db 30 CTTTGTGCTTATGGTATCCAGGATTAGAGAAAGCCATTTCTGGTTGGCTTCCCT 89

QY 123 GTGTGGCTTAGGCAGACGCCACAATTTGGGCAATATACTATCTGGTTGGTTGGCCAC 182
Db 90 CTTTCCATGTATGTAGTGGCAATGTTTGGAAATGCACTGGTGGTCTTCATCGTAAGGAC 149

QY 183 TGAACAGCTTGGCAAGCCTGTGTACCTTTTCTGTGCATGCTCTCAACCATCGACTT 242
Db 150 GGAACGACCTGACGCTCGGATGACCTTTTCTGTGCATGCTTGCAGCCATTTGACCT 209

QY 243 GGCTGCTCTGTCTCCACAGTTCCTCAAGCTACTGGCTATCTTCTGGTGTGGAGCCGACA 302
Db 210 GGCTTATCCACATCCACCATGCTTAAGATCTTGTGCTTCTTCTGGTTGATTCCCGAGA 269

QY 303 TATATCTGCTCTGCTGCTGCTGCGCAATATGTTCTTCTATTCATGCTCTCTGCATGATGA 362
Db 270 GATTAGCTTTAGGCTGTCTTACCAGATGTTCTTATTCATGCTCTCTCAGCCATTGA 329

QY 363 GTCCATGTCTACTGGCCATGCTTGTATCGCTAGCTGGCCATCTGCCACCATCGG 422
Db 330 ATCCACCATCTCTGCTGGCCATGCTTGTACCGTTATGTGGCCATCTGCCACCATCGG 389

QY 423 CTATGCCACAATCTCTCACTGACACCATCATTTGCCACATAGGGGTGGCAGCTGTAGTGGG 482
Db 390 CCATGCTGAGTCTCAACATACAGTACAGCCAGCCAGATGGCATCTGGCTGTGGTCCG 449

RESULT 4
US-08-465-980-1
; Sequence 1, Application US/08465980
; Patent No. 5756309
; GENERAL INFORMATION:
; APPLICANT: Soppet, Daniel R.
; APPLICANT: Li, Yi
; APPLICANT: Rosen, Craig A.
; APPLICANT: Ruben, Steven M.
; TITLE OF INVENTION: HUMAN G-PROTEIN RECEPTOR HPRAJ70
; NUMBER OF SEQUENCES: 8
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: CABELLA, BYRNE, BAIN, GILFILLAN, CECCHI,
; ADDRESSEE: STUART & OLSTEIN
; STREET: 6 Becker Farm Road
; CITY: Roseland
; STATE: New Jersey
; COUNTRY: USA
; ZIP: 07068
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/465,980
; FILING DATE: 06-JUN-1995
; CLASSIFICATION: 536
; ATTORNEY/AGENT INFORMATION:
; NAME: Ferraro, Gregory D.
; REGISTRATION NUMBER: 36,134
; REFERENCE/DOCKET NUMBER: 325800-446
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 201-994-1700
; TELEFAX: 201-994-1744
; INFORMATION FOR SEQ ID NO: 1:

Qy	243	GGCGCCTCTGTCTCCACAGTTCCCAAGCTACTGGGCTATCTTCTGTGGTGTGGAGCCGGACA	302
Db	483	GGCCTTATCCACATCCACCATGCGCTAAGATCCTTGCCCTTTCTGTGTTTGAATCCCGAGA	542
Qy	303	TATATCGCCTCTGCCTGCGCTGGACATATGTTCTTCAATCATGCCCTTCGTCATGATGGA	362
Db	543	GATTAGCAATTGAGGCCCTGTCTTACCCAGATGTTCTTATTATCATGCCCTCTCAGCCATGGA	602
Qy	363	GTCACCTGTGCTACTGCGCCATGGCCTTTGATCGCTAGCTGGCCATCTGCCACCCACTCCG	422
Db	603	ATCCACCATCTGCTGGCCATGGCCTTTGACCGTTATGTGGCCATCTGCCACCCACTGCG	662
Qy	423	CTATGCCACAATCTCTACTGACACATCATATTGCCCAATAGGGGTGGAGCTGTAGTGGG	483
Db	663	CCATGCTGCAGTGTCTAACAAATACAGTAACAGCCACAGATTGGCATCGTGGCTGTGGTCCG	722
Qy	483	AGGCTCCCTGCTCATGCTCCCATGTCCTTCTTTATTGGGCGTTTGAACCTTCTGCCAAG	542
Db	723	CGGATCCCTCTTTTTTTCACATGCGCTCTGCTGATCAAGGGCTGGCCTTCTGCCACTC	782
Qy	543	CCATGTGATCCTACACACGTACTGTGAGCACATGGCTGTGTGAAGCTGGCCTGTGGAGA	602
Db	783	CAATGTCTCTCGCACTCTTATTGTGTCCACAGAGTGAATGAAGTTGGCCTATGCAGA	842
Qy	603	CACAGGCGCTAACCGTGTGTATGGGTGACAGCTGCACTGTTGTGTCATTTGGGGTTGACTT	662
Db	843	CACTTTGGCCCAATGTGGTATATGGTCTTACTTGCCATTCTGCTGGTTCATGGGCGTGGACGT	902
Qy	663	GTTTTGCAATTGGTCTCTCCTATGCCCTAATTGCAACAGCTCTCCTTCGCCCTCATCCCA	722
Db	903	AATGTTTCATCTCCTGTGCTCTATTTTCTGATTAATACGAACGGTTCTGCAACTGCCTTCCAA	962
Qy	723	TGAAGCTCGGTCCAAAGGCCCTAGGAGCCTGTGGTTCCCATCTCTGTGTCACTCATCTC	782
Db	963	GTCAAGGGGCGCAAGGCCCTTTGNAACCTGTGTGTCAACATTTGGTGTGTACTCGCCTT	1022
Qy	783	TTATACACAGCCCTCTTCTCTCTTTTATACACGCTTTTGGCCATCAAGTTCCAGTCCA	842
Db	1023	CTATGTGCCACTTATTGGCCTCTCAGTTTGTACACCGCTTTTGAAACAGCCTTCATCCCAT	1082
Qy	843	TATTCACTTCTTTTGGCCAAATGTTATCTGCTTTTGCCACTGTCTTATCTCTGTGGT	902
Db	1083	TGTGCGTGTGTGATGGGTGCATCTACCTGCTGCTGCTCTCTGTGCATCAATCCCATCAT	1142
Qy	903	ATATGGAGTTAAGACCAACACAGATCCGTAAGAGTTGTCAAGGTGTTTCAAGTGGCA	962
Db	1143	CTATGGTGGCAAAACCAACACAGATCAGAACCGGTGCTGGCTATGTTCAAGATCAGCTG	1202
Qy	963	GGGAATGGGCATCAAGC	980
Db	1203	TGACAAGGACTTGCAGGC	1220

RESULT 6

```

US-09-339-115-1
; Sequence 1, Application US/09339115
; Patent No. 6372891
; GENERAL INFORMATION:
; APPLICANT: Soppet, Daniel R.
; APPLICANT: Li, Yi
; APPLICANT: Rosen, Craig A.
; APPLICANT: Ruben, Steven M.
; TITLE OF INVENTION: HUMAN G-PROTEIN RECEPTOR HPAJ70
; NUMBER OF SEQUENCES: 8
; CORRESPONDENCE ADDRESS:
; ADDRESS: CARELLA, BYRNE, BAIN, GILFILLAN, CECCHI
; ADDRESSEE: STUART & OLSTEIN
; STREET: 6 Becker Farm Road
; CITY: Roseland
; STATE: New Jersey
; COUNTRY: USA
; ZIP: 07068
; COMPUTER READABLE FORM:

```

QY 663 GTTTTCATGCTCTCTCTATGCCCTAAATGACAAAGCTGCTCTCGCTCTCATCCCA 722
Db |||||
QY 903 AATGTTCACTCTCTCTCTATTTCTGATTAATACGAACGGTTCTGCAACTGCTTCCAA 962
Db |||||
QY 723 TGAAGCTCGGTCAAGGCCCTAGGACCTGTGTTTCCCATGTCTGTGTCTATCTCATCTC 782
Db |||||
QY 963 GTCAAGCGGGCCAAAGCCTTTGGAACCTGTGTGTACACATTTGGTGTGTAATCGCCTT 1022
QY 783 TTATACACAGCCCTCTCTCTCTTTTACACACCGCTTTGGCCATCACGTTCCAGTCCA 842
Db |||||
QY 1023 CTATGTGCCACTTATTTGGCCTCTCAGTTGTACACCGCTTTGGAACAGCCCTTCATCCCA 1082
QY 843 TATTACACATTTCTTTGGCCAAATGTTTATCTGCTTTTGGCCACCTGCTCTTAATCCTGTGGT 902
Db |||||
QY 1083 TGTGCTGTGTGTCATGGGTGACATCTACTGCTGTGCTCTCTGTCATCAATCCCAATCAT 1142
QY 903 ATATGGAGTTAAGACCAACAGATCCGTAAAGAGTTGTACAGGTGTTTCAAAGTGGGCA 962
Db |||||
QY 1143 CTATGGTGCACAAACCAACAGATCAGAACACGGGTGCTGGCTATGTTCAAGATCAGCTG 1202
QY 963 GGAATGGGCATCAAGGC 980
Db |||||
QY 1203 TGACAAGGACTTGCAGGC 1220

RESULT 7

PCT-US95-07093-1
; Sequence 1, Application PC/TUS9507093
; GENERAL INFORMATION:
; APPLICANT: Soppet, Daniel R.
; APPLICANT: Li, Yi
; APPLICANT: Rosen, Craig A.
; APPLICANT: Ruben, Steven M.
; TITLE OF INVENTION: HUMAN G-PROTEIN RECEPTOR HPRAJ70
; NUMBER OF SEQUENCES: 8
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: CARELLA, BYRNE, BAIN, GILFILLAN, CECCHI,
; ADDRESS: STUART & OLSTEIN
; STREET: 6 Becker Farm Road
; CITY: Roseland
; STATE: New Jersey
; COUNTRY: USA
; ZIP: 07068
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US95/07093
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/465,980
; FILING DATE: 06-JUN-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Ferraro, Gregory D.
; REGISTRATION NUMBER: 36,134
; REFERENCE/DOCKET NUMBER: 325800-446
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 201-994-1700
; TELEFAX: 201-994-1744
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1474 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 274..1233

PCT-US95-07093-1

Query Match 33.5%; Score 330.8; DB 5; Length 1474;
Best Local Similarity 60.0%; Pred. No. 2.4e-98;
Matches 551; Conservative 0; Mismatches 367; Indels 0; Gaps 0;
QY 63 CTTCTCTCTCTGGGCAATCCAGGCTCTGGAACAATTTCAATTTGGTCTCTCACTCCCTGT 122
Db CTGTGTGCTTATTTGTTATCCCAAGATTAGAGAAGCCCAATTTCTGGGTTGGCTTCCCTCT 362
QY 123 GTGTGGCTTAGGCACAGCCACAATTTGGGCAATATACTATTCTGTTGTTGTTGCCAC 182
Db CTTTTCATGATGATAGTGGCAATGTGTGGAACCTGTGCTGTTGTTCTTCAATCGTAAGGAC 422
QY 183 TGAACCAAGCTTTGCAACAAGCTGTGTACCTTTTCTGTGTCATGCTCTCAACCATGACATT 242
Db GGACGGCAGCTGTCAGCTCCGATGATCTCTTTCTGTGATGCTGTGACGCAATGACCT 482
QY 243 GGCTGCTCTGTCTCCACAGTTCCCAAGTACTGCTATCTTCTGGTGTGGAGCCGGACA 302
Db GGCCTTATCCACATCCACATGCTTAAAGATCTTCCCTTTTCTGGTTGATTTCCCGAGA 542
QY 303 TATATCTGCTCTGCTGCTGGCACATATGTTCTTCAATTCATGCTCTTCGATGATGA 362
Db GATTAGCATTTGAGGCTCTCTTACCCAGATGTTCTTTATTCATGCCCTCTAGGCCATTGA 602
QY 363 GTCCACTGTGCTATGCGGCATGGCCTTTGATGCTGTACGTGGCCATCTGCCACCCACTCG 422
Db ATCCACCATCTGCTGGCCATGGCCTTTGACCGTTATGTGGCCATCTGCCACCCACTCG 662
QY 423 CTATGCCACAATCCTCACTGACACATCATTTGCCACATAGGGGTGGCAGCTGTAGTGG 482
Db CCAATGCTCAGTGTCTCAACAATACAGTAACAGCCAGATTTGGCACTGCTGGTGTGGTCCG 722
QY 483 AGGCTCCCTGCTCAATGCTCCCATGCTCTTTTATTTGGGGTGTGAACTTCTGCAAAAG 542
Db CGGATCCCTCTTTTTCCTCCACTGCTCTGATCAAGGGCTGGGCTTCTGCACTC 782
QY 543 CCATGTGATCTTACACAGTACTGTGAGCAATGCTGTGTGAGCTGGCTGTGGAGA 602
Db CAATGCTCTCTCGCACTCTTATTTGTCTCCACAGGATGTAATGAAGTTGGCTATGAGA 842
QY 603 CACGAGGCTAAACCGTGTGATGGCTGACAGCTGCACTGTGTGTCATTTGGGTTGACTT 862
Db CACTTTGCCCAATGTGTATATGTTCTTACTGCAATCTCTGCTGTGTCATGGGCGTGGAGT 902
QY 663 GTTTTGCATTTGCTCTCTCTATGCTTAAATGCAACAGCTGCTTGGCTCTCATCCCA 722
Db AATGTTCACTCTCTTGTCTTATTTCTGATAATAGAACGTTCTGCAACTGCTTCCAA 962
QY 723 TGAAGCTCGGTCCAAAGGCCCTAGGACCTGTGGTTCCCATGTCTGTGTCTCTCATCTC 782
Db GTCAAGCGGGCCAAAGCCTTTGGAACCTGTGTGTCAACATTTGGTGTGTTACTGCTT 1022
QY 783 TTATACACAGCCCTCTCTCTCTTTTACACACCGCTTTGGCCATCACGTTCCAGTCCA 842
Db CTATGTGCCACTTATTTGGCCTCTCAGTTGTACACCGCTTTTGGAAACAGCCCTTCATCCAT 1082
QY 843 TATTACATTTCTTTGGCCAAATGTTTATCTGCTTTTGGCCACCTGCTCTTAATCCTGTGGT 902
Db TGTGCTGTTGTGTCATGGGTGACATCTACTGCTGCTGCTCTCTCTGTCATCAATCCATCAT 1142
QY 903 ATATGGAGTTAAGACCAACAGATCCGTAAAGAGTTGTGAGGGTGTTCAGAGTGGGCA 962
Db CTATGGTGCACAAACCAACAGATCAGAACACGGGTGCTGGCTATGTTCAAGATCAGCTG 1202
QY 963 GGAATGGGCATCAAGGC 980
Db |||||
QY 1203 TGACAAGGACTTGCAGGC 1220

RESULT 8

US-09-668-680-13

Db	1177	ATTGTCTATGGGCTAACTGTGGCTCTGTCTGGGCATCGGACTGGATTCGATTCCTCATTTGCC	1233
QY	676	CTCTCTATGCCCTAATTGCAAGCTGTCCTTCGGCTCTCATCCCATGAAGCTCGGTCC	735
Db	1237	ATTTCCTATGGGTTTATCTCCATGACAGTCTTTACCTTCCATCTCATGATGCCAGCAC	1296
QY	736	AAGGCCCTAGGACCTGTGGTTCCCATGTCTGTGCATCCTCATCTCTTATACACCAGCC	795
Db	1297	AAAGCTCTAGTACCTGTGGCTCCACATATGGCATCATCTGGTTTCTACATCCCTGCC	1356
QY	796	CTCTCTCTCTTTTATACACACCGCTTTGGCCATCAC---GTTCCAGTCCATATTCACATT	852
Db	1357	TTCTCTCTCTCTCCACCACCGCTTTGGTCACACGAAGTCCCAAGCATGTGCACATC	1416
QY	853	CTTTTGGCCAAATGTTTATCTGCTTTTGGCACCTGCTCTTAATCCTCTGGGTATATGGAGTT	912
Db	1417	TTTCTGGCTAAATCTCATGTGTGGTGGCTCTGTACTCAATCCTATTTCTCTATGGAGCT	1476
QY	913	AAGACCAAAACAGATCCGTAACAGAGTTGTCCAGGCTGTTTCAAAAGTGGGCAG	963
Db	1477	AGAACCAAGAGATTCGGAGTCGACTTCTAAACTGCTTCACTCCCTGGGAG	1527
RESULT 9			
US-09-016-434-1312			
; Sequence 1312, Application US/09016434			
; Patent No. 6500938			
; GENERAL INFORMATION:			
; APPLICANT: Janice Au-Young			
; APPLICANT: Jeffrey J. Seilhamer			
; TITLE OF INVENTION: COMPOSITION FOR THE DETECTION OF SIGNALING			
; TITLE OF INVENTION: PATHWAY GENE EXPRESSION			
; NUMBER OF SEQUENCES: 1490			
; CORRESPONDENCE ADDRESS:			
; ADDRESSEE: INCYTE PHARMACEUTICALS, INC.			
; STREET: 3174 PORTER DRIVE			
; CITY: PALO ALTO			
; STATE: CALIFORNIA			
; COUNTRY: USA			
; ZIP: 94304			
; COMPUTER READABLE FORM:			
; MEDIUM TYPE: Floppy disk			
; COMPUTER: IBM PC compatible			
; OPERATING SYSTEM: PC-DOS/MS-DOS			
; SOFTWARE: Word Perfect 6.1 for Windows/MS-DOS 6.2			
; CURRENT APPLICATION DATA:			
; APPLICATION NUMBER: US/09/016,434			
; FILING DATE: HERewith			
; CLASSIFICATION:			
; PRIOR APPLICATION DATA:			
; APPLICATION NUMBER:			
; FILING DATE:			
; CLASSIFICATION:			
; ATTORNEY/AGENT INFORMATION:			
; NAME: Zeller, Karen J.			
; REGISTRATION NUMBER: 37,071			
; REFERENCE/DOCKET NUMBER: PA-0002 US			
; TELECOMMUNICATION INFORMATION:			
; TELEPHONE: (650) 855-0555			
; TELEFAX: (650) 845-4166			
; INFORMATION FOR SEQ ID NO: 1312:			
; SEQUENCE CHARACTERISTICS:			
; LENGTH: 1854 base pairs			
; TYPE: nucleic acid			
; STRANDEDNESS: single			
; TOPOLOGY: linear			
; IMMEDIATE SOURCE:			
; LIBRARY: GENBANK			
; CLONE: G32085			
US-09-016-434-1312			
Query Match		14.4%;	Score 142.4; DB 4; Length 1854;

TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
US-08-748-506-6

Query Match 13.4%; Score 131.8; DB 3; Length 966;
Best Local Similarity 49.8%; Pred. No. 5.3e-33;
Matches 328; Conservative 0; Mismatches 330; Indels 0; Gaps 0;

QY 170 TTGTTGTTGACCACTGAAACAGCTCTTGGACAGCCTGTGTACCTTTTCTGTGCACTGCTCT 229
Db 158 TTGCTGTTGTTGACCACTGAAACAGCTCTTGGACAGCCTGTGTACCTTTTCTGTGCACTGCTCT 217

QY 230 CAACCATCGACTGTGCTGCTGTCTTCCACAGTTTCCCAAGTACTGCTATCTTCTGCT 289
Db 218 CTCTCTGGAGATTGGCTATATCTTCTGTCTGTATACCAAGATGCTGAGAGCCTTGTGA 277

QY 290 GTGAGCCGGACATATATCTGCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 349
Db 278 GTGAGCCGGACATATATCTGCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 337

QY 350 TCTGCTGATGAGTCCACTGCTGTCTTCCACAGTTTCCCAAGTACTGCTATCTTCTGCT 409
Db 338 TTGCTGATGAGTCCACTGCTGTCTTCCACAGTTTCCCAAGTACTGCTATCTTCTGCT 397

QY 410 GCCACCCACTCCGCTATGCCAAATCTCTCACTGACACCATCATTTGCCACATAGGGGTGG 469
Db 398 GCTCCCACTCCACTATGCAACCCGATGAGTCTGAGGTATGTGCCATTTGGCAATGG 457

QY 470 CAGCTGATGCGAGGCTCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 529
Db 458 TTTTCATGGGAATGGGATGCTAGTGTCTGCGACAGACCAATTTNATTTNCTCTTGA 517

QY 530 ACTTCTGCCAAGCCATGCTGATCTTACACAGTACTGTGACACATGCTGCTGCTGCTGCT 589
Db 518 ACTTCTGCGACCTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 577

QY 590 TGGCCTGTGGAGACACAGGCTTAAACGCTGTGTATGGGCTGACAGCTGCACTGTTGTCA 649
Db 578 TTGCTGTGTTGATATATCTTCCAAAGCGGCTGCCATCTTTGTGCGACCAATCTCTGTA 637

QY 650 TTGGGTTGATGTTTGTGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 709
Db 638 TATCTAGTCCATTTTGTGATCTTTTATTTTATGTCAGAAATCTGCTGCTGCTGCTGCT 697

QY 710 GCCTCTCATCCCATGAAGCTCGGTCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 769
Db 698 TGATGCTTACCTGAGGGGGCCATAGCTCTTCCACCTGTTCTCACCCTACTTGT 757

QY 770 TCATCTCATCTCTTATACACAGCCTCTTCTCTCTTTTATACACAGCCTTTTGGCCA 827
Db 758 TAGTCACACTCTTTTATGCTGTGCTTACCTTATTTGAGGCCCAAGTCTAGCCA 815

RESULT 13
US-08-467-948A-1
Sequence 1, Application US/08467948A
Patent No. 5998164
GENERAL INFORMATION:
APPLICANT: LI, YI
APPLICANT: CAO, LIANG
APPLICANT: NI, JIAN
APPLICANT: GENTZ, REINER
APPLICANT: BULT, CAROL J.
APPLICANT: SUTTON III, GRANGER G.
APPLICANT: ROSEN, CRAIG A.
TITLE OF INVENTION: Polynucleotides Encoding Human G-Protein
TITLE OF INVENTION: Coupled Receptor GPR
NUMBER OF SEQUENCES: 30
CORRESPONDENCE ADDRESS:
ADDRESSEE: STERN, KESSLER, GOLDSTEIN & FOX P.L.L.C.
STREET: 1100 NEW YORK AVE., NW, SUITE 600
CITY: WASHINGTON
STATE: DC

COUNTRY: USA
ZIP: 20005
COMPUTER READABLE FORM:
MEDIUM TYPE: FLOPPY DISK
COMPUTER: IBM PC COMPATIBLE
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PATENTIN RELEASE #1.0, VERSION #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/467,948A
FILING DATE: 06-JUN-1995
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PCT/US95/04079
FILING DATE: 30-MAR-1995
ATTORNEY/AGENT INFORMATION:
NAME: STEFFE, ERIC K.
REGISTRATION NUMBER: 36,688
REFERENCE/DOCKET NUMBER: 1488.1140003/EKS/KLM
TELECOMMUNICATION INFORMATION:
TELEPHONE: 202-371-2600
TELEFAX: 202-371-2540
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 1713 base pairs
TYPE: nucleic acid
STRANDEDNESS: both
TOPOLOGY: both
MOLECULE TYPE: cDNA
FEATURE:
NAME/KEY: CDS
LOCATION: 116..1003
US-08-467-948A-1

Query Match 13.3%; Score 131; DB 2; Length 1713;
Best Local Similarity 48.1%; Pred. No. 1.4e-32;
Matches 371; Conservative 0; Mismatches 400; Indels 0; Gaps 0;

QY 59 CTGCT 118
Db 153 CAGAGTTCT 212

QY 119 CTGCTGTGCTTCT 178
Db 213 TCTTCT 272

QY 179 CCACTGAACCACT 238
Db 273 CACTGGAATCCAGACTCCACACCCCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCA 332

QY 239 ACTTGGCTGCTCTGCT 298
Db 333 ACATCGCTATGCTCTGCTCA 392

QY 299 GACATATATCTGCTCA 358
Db 393 AGCCCATCA 452

QY 359 TGGAGTCCACTGTGCT 418
Db 453 CTGAATGCT 512

QY 419 TCCGCTATGCCACAATCT 478
Db 513 TCCGATATTTCT 572

QY 479 TGGAGGCT 538
Db 573 CATGTGCT 632

QY 539 AAAGCCATGTGATCT 598
Db 633 GGCTCTGTGAATCAACCACT 692

QY 599 GAGACACAGCCCTAACGGTGTGTATGGGCTGACAGCTGCACTGTTGGTCAATGGGGTGG 658
Db 693 CTGATACCTGGCTCAACAGGTGTGTATCTTTGAAAGCTGCAATGTTATCTGGTGGGAC 752
QY 659 ACTTGTGTTGATGGTCTCTCTATGCGCCCTAAATGACAAAGCTGCTCTGCGCTCTCAT 718
Db 753 CACTCTGCTGGTGTCTCTCTACTACACATCTCTGGGGGATCCTGAGGATCCAGT 812
QY 719 CCCATGAAGTCTGGTCCAAAGCCCTAGGACCTGTGGTTCCCATGTCTGTGTCAATCTCA 778
Db 813 CTGGGAGGGCCGAGAAAGCCCTTCTCCACCTGCTCTCCACCTCTGCGTAGTGGGAC 872
QY 779 TCTCTTATACACAGCCCTCTCTCTCTCTTTTATACACAGCCGTTTGGGCATC 829
Db 873 TCTCTTTTGGAGCCCATCGTATGATGAGTGGCCCTTAAGTCCCGGCATC 923

RESULT 14

US-08-467-947A-1
; Sequence 1, Application US/08467947A
; Patent No. 6090575
; GENERAL INFORMATION:
; APPLICANT: LI, YI
; APPLICANT: CAO, LIJANG
; APPLICANT: NI, JIAN
; APPLICANT: GENTZ, REINER
; APPLICANT: BULT, CAROL J.
; APPLICANT: SUTTON III, GRANGER G.
; APPLICANT: ROSEN, CRAIG A.
; TITLE OF INVENTION: Polynucleotides Encoding Human G-Protein
; TITLE OF INVENTION: Coupled Receptor GPR1
; NUMBER OF SEQUENCES: 30
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: STERNE, KESSLER, GOLDSTEIN & FOX P.L.L.C.
; STREET: 1100 NEW YORK AVE., NW, SUITE 600
; CITY: WASHINGTON
; STATE: DC
; COUNTRY: USA
; ZIP: 20005

COMPUTER READABLE FORM:
; MEDIUM TYPE: FLOPPY DISK
; COMPUTER: IBM PC COMPATIBLE
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PATENTIN RELEASE #1.0, VERSION #1.30

CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/467,947A

FILING DATE: 06-JUN-1995

CLASSIFICATION: 435

PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/US95/04079

FILING DATE: 30-MAR-1995

ATTORNEY/AGENT INFORMATION:
; NAME: STEFFE, ERIC K.

REGISTRATION NUMBER: 36,688

REFERENCE/DOCKET NUMBER: 1488.1140002/EKS/KLM

TELECOMMUNICATION INFORMATION:
; TELEPHONE: 202-371-2600

TELEFAX: 202-371-2540

INFORMATION FOR SEQ ID NO: 1:

SEQUENCE CHARACTERISTICS:

LENGTH: 1713 base pairs

TYPE: nucleic acid

STRANDEDNESS: both

TOPOLOGY: both

MOLECULE TYPE: cDNA

FEATURE:
; NAME/KEY: CDS

LOCATION: 116..1003

US-08-467-947A-1

Query Match 13.3%; Score 131; DB 3; Length 1713;
Best Local Similarity 48.1%; Pred No. 1.4e-32;
Matches 371; Conservative 0; Mismatches 400; Indels 0; Gaps 0;

QY 59 CTGTCCTTCTCTCTCTGGGATCCAGGCTGGGACAAATTTCAATTTTGGTCTCTCATCC 118
Db 153 CAGAGTTCTCTCTCTCTGGGATTTCTCTGGGCCCCAAGGATTCAGATGCTCTCTTGGGC 212
QY 119 CTGTGTCTGGCTTAGGCACAGCCACAATTTGGGCAATATAAATATTTCTGTGTTGTTG 178
Db 213 TCTTCT 272
QY 179 CCACTGAACAGTCTTTCACAAAGCCTGTGACCTTTTCTGTGTGATGCTCTCTCAACCAT 238
Db 273 CACTGGACTCCAGACTCCACACCCCATGTACTTCTCTCTCTCTCTCTCTCTCTCTCTCT 332
QY 239 ACTTGGCTGCTCTGTCTCCACAGTTCCTCAAGCTACTGGCTATCTTCTGTGTGTGGAGCG 298
Db 333 ACATCGGCTATGCTGCAACACAGTGGCCAGATGTGGTGAACCTCTCTGATCCAGCA 392
QY 299 GACATATATCTGCTCTGCTGCTGGCAGATATGTTCTTCTCATTCATGCTCTCTGATGA 358
Db 393 AGCCATCTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 452
QY 359 TGGAGTCCACTGTGCTACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 418
Db 453 CTGAATGCTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 512
QY 419 TCGCTATGCCAATCTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 478
Db 513 TCCGATATTTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 572
QY 479 TGGAGGCTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 538
Db 573 CATGTGCTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 632
QY 539 AAAGCCATGTGATCTCTACACAGTCTGTGTGAGCAGATGGTGTGTGTGTGTGTGTGTGT 598
Db 633 GGCCTCGTGAATCAACCACTTCTCTGTGAAATCTCTGCTCTCTCTCTCTCTCTCTCTCT 692
QY 599 GAGACACAGCCCTAACGGTGTGTATGGGCTGACAGCTGCACTGTTGGTCAATGGGGTGG 658
Db 693 CTGATACCTGGCTCAACAGGTGTGTATCTTTTGAAGCTGATGTTCTCTCTCTCTCTCTCT 752
QY 659 ACTTGTCTTCT 718
Db 753 CACTCTGCTGCTGCTGCTGCT 812
QY 719 CCCATGAAGTCTGGTCCAAAGCCCTAGGACCTGTGGTTCCCATGTCTGTGTCTCATCTCA 778
Db 813 CTGGGAGGGCCGAGAAAGCCCTTCTCCACCTGCTCTCTCTCTCTCTCTCTCTCTCTCTCT 872
QY 779 TCTCTTATACACAGCCCT 829
Db 873 TCTTCTTTTGGAGCCCATCGTATGATGAGTGGCCCTTAAGTCCCGGCATC 923

RESULT 15

US-08-748-506-8
; Sequence 8, Application US/08748506
; Patent No. 6159707
; GENERAL INFORMATION:
; APPLICANT: Ronnett et al.
; TITLE OF INVENTION: NOVEL SPERM RECEPTORS
; NUMBER OF SEQUENCES: 31
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Leydig, Voit & Mayer, Ltd.
; STREET: Two Prudential Plaza, Suite 4900
; CITY: Chicago
; STATE: IL
; COUNTRY: US
; ZIP: 60601-6780
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/748,506
FILING DATE: 08-NOV-1996
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 60/033,751
FILING DATE: 09-NOV-1995
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
REFERENCE/DOCKET NUMBER: 74940
TELEPHONE: 312-616-5600
TELEFAX: 312-616-5700
INFORMATION FOR SEQ ID NO: 8:
SEQUENCE CHARACTERISTICS:
LENGTH: 966 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
US-08-748-506-8

Job time : 96 secs

Query Match 12.8%; Score 126.8; DB 3; Length 966;
Best Local Similarity 49.5%; Pred. No. 2.3e-31;
Matches 326; Conservative 0; Mismatches 332; Indels 0; Gaps 0;

QY	170	TTGTTGTTGGCCACTGAACCACTCTTGGACAGCCCTGCTGCTACCTTTTCTGTGCATGCTCT	229
Db	158	TGCTATTGTTACCACTCTCTACACACCCCATGCTCTCTTTCTGGCCAACTTGT	217
QY	230	CAACCATCGACTTGGCTGCCCTCTGTCTCCACAGTTCCCAAGCTACTGCTATCTTCGGT	289
Db	218	CTCTCTGGAGATTGGCTATCTTCTCTGTCTATACCCAGATGCTGCAGAGCTTTGGA	277
QY	290	GTGAGCGCGCATATATCTGCCCTCTGCCCTGGGCACATATGTTCTTCATTTCATGCCCT	349
Db	278	GTGAGGCCCGAGAGATCTTTCAGGTGGGATGTGCCACACAGATGTTTTCTTCATATTCT	337
QY	350	TCTGCATGATGGAGTCCACTGTCTGTCTACTGGCCATGGCCTTTGATCGCTACGTGGCCATCT	409
Db	338	TTGTTATTAATGATGTGCTCTATTTGGCAGCCATGGCCCTTTGACCGCTATATGGCTATAT	397
QY	410	GCCACCCACTCCGCTATGCCAAATCTCTCACTGACACACCATCATTTGCCACATAGGGTGG	469
Db	398	GTTCCCACTCCACTATGCAACCCGAATGAGTCTGTAGGTATGTGCCACTTGGCAATTG	457
QY	470	CAGCTGTAGTCCGAGGCTCCCTGTCTCATGTCCCATGTCCTTCTTTATTTGGGCGTTGA	529
Db	458	TTTCATGGGTGATGGGATGCATAGGTCTGGGACACAGCAATTTTATTTCTCTCTGA	517
QY	530	ACTTCTGCCAAGCCATGTGATCTCTACACAGTACTGTGACACATGGCTGTGGTGAAGC	589
Db	518	ACTTCTGTGGACCTGTGAGATAGACCACTTCTCTGTGATCTTCCACCTCTCTGGCAC	577
QY	590	TGGCCTGTGAGACACACAGGCTTAACCGTGTGTATGGGCTGACAGCTGCATGTTGTCA	649
Db	578	TTGCCTGTGGTGATACATCCAAATTGAGGCTGCCATCTTTGTGTAGTTGTCTCTGCA	637
QY	650	TTGGGGTTGACTGTTTGTGATTTGGTCTCTCTATAGCCCTTAATTGACAGAGCTGTCTTC	709
Db	638	TATCTAGCCCTTTTCTGCTGATCATTTTATTCTTATGTAGAAATCTCGTTGCAGTGCTGG	697
QY	710	GCCTCTCATCCCATAGCTCGGTCCCAAGGCCCTTAGGGACCTGTGGTTCCTCATGCTGTG	769
Db	698	TGATGCTTTCACCTGAGGGGGCCACAAAGCCCTTCAACTGTCTCTCCCACTACTTG	757
QY	770	TCATCTCATCTCTTATACACAGCCCTCTTCTCTCTTTTATACACACCGCTTTGGCCA	827
Db	758	TAGTCACACTCTTTTATGGCTCAGGATCTGTACCTATTGAGGCGCTAAGTCTAGCCA	815

Search completed: August 27, 2004, 18:30:17

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 27, 2004, 16:19:27 ; Search time 540 Seconds

(without alignments)
8995.273 Million cell updates/sec

Title: US-10-081-775-1_COPY_537_1523

Perfect score: 987

Sequence: 1 atgtccagcactcttgcca.....tgggcatcaagcattctgag 987

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 3237270 seqs, 2460713050 residues

Total number of hits satisfying chosen parameters: 6474540

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications NA:*

1: /cgn2_6/ptodata/2/pubpna/US07_PUBCOMB.seq:
2: /cgn2_6/ptodata/2/pubpna/PCT_NEW_PUB.seq:
3: /cgn2_6/ptodata/2/pubpna/US05_NEW_PUB.seq:
4: /cgn2_6/ptodata/2/pubpna/US06_PUBCOMB.seq:
5: /cgn2_6/ptodata/2/pubpna/US07_NEW_PUB.seq:
6: /cgn2_6/ptodata/2/pubpna/PCTUS_PUBCOMB.seq:
7: /cgn2_6/ptodata/2/pubpna/US08_NEW_PUB.seq:
8: /cgn2_6/ptodata/2/pubpna/US09_PUBCOMB.seq:
9: /cgn2_6/ptodata/2/pubpna/US09A_PUBCOMB.seq:
10: /cgn2_6/ptodata/2/pubpna/US09B_PUBCOMB.seq:
11: /cgn2_6/ptodata/2/pubpna/US09C_PUBCOMB.seq:
12: /cgn2_6/ptodata/2/pubpna/US09_NEW_PUB.seq:
13: /cgn2_6/ptodata/2/pubpna/US10_NEW_PUB.seq:
14: /cgn2_6/ptodata/2/pubpna/US10A_PUBCOMB.seq:
15: /cgn2_6/ptodata/2/pubpna/US10B_PUBCOMB.seq:
16: /cgn2_6/ptodata/2/pubpna/US10C_PUBCOMB.seq:
17: /cgn2_6/ptodata/2/pubpna/US10_NEW_PUB.seq:
18: /cgn2_6/ptodata/2/pubpna/US60_NEW_PUB.seq:
19: /cgn2_6/ptodata/2/pubpna/US60_PUBCOMB.seq:

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	987	100.0	1567	15	US-10-081-775-1
2	980.6	99.4	990	9	US-09-886-055-262
3	980.6	99.4	990	10	US-09-804-291-262
4	980.6	99.4	990	13	US-10-343-650A-353
5	980.6	99.4	998	15	US-10-025-806-35
6	980.6	99.4	998	15	US-10-025-806-37
7	980.6	99.4	1366	15	US-10-017-161-923
8	980.6	99.4	1390	16	US-10-292-798-797
9	956.6	96.9	966	17	US-10-297-021-42
10	376.2	38.1	980	11	US-09-844-861A-25
11	376.2	38.1	1345	15	US-10-017-161-879
12	376.2	38.1	1986	9	US-09-864-761-3771
13	375.2	38.0	945	9	US-09-886-055-90
14	375.2	38.0	945	10	US-09-804-291-90

15	375.2	38.0	945	13	US-10-182-822A-25	Sequence 25, Appl
16	375.2	38.0	945	13	US-10-343-650A-161	Sequence 161, App
17	375.2	38.0	945	14	US-10-032-106-5	Sequence 5, Appli
18	374.6	38.0	1345	16	US-10-292-798-757	Sequence 757, App
19	373.6	37.9	945	16	US-10-387-629-161	Sequence 161, App
20	369.8	37.5	945	9	US-09-864-761-20337	Sequence 20537, A
21	368.4	37.3	1013	15	US-10-025-806-25	Sequence 25, Appl
22	368.4	37.3	1360	15	US-10-017-161-947	Sequence 947, App
23	368.4	37.3	1360	16	US-10-292-798-821	Sequence 821, App
24	367.2	37.2	982	11	US-09-844-861A-21	Sequence 21, Appl
25	366.6	37.1	1345	15	US-10-017-161-877	Sequence 877, App
26	365	37.0	1345	16	US-10-292-798-755	Sequence 755, App
27	364	36.9	945	16	US-10-387-629-163	Sequence 163, App
28	364	36.9	1151	11	US-09-844-861A-19	Sequence 19, Appl
29	363.4	36.8	1400	9	US-09-886-055-86	Sequence 86, Appl
30	362.4	36.7	1400	10	US-09-804-291-86	Sequence 86, Appl
31	362.4	36.7	945	13	US-10-343-650A-159	Sequence 159, App
32	356.6	36.1	960	9	US-09-886-055-96	Sequence 96, Appl
33	356.6	36.1	960	10	US-09-804-291-96	Sequence 96, Appl
34	356.6	36.1	960	13	US-10-343-650A-317	Sequence 317, App
35	356.6	36.1	960	16	US-10-387-629-159	Sequence 159, App
36	355.6	36.0	990	13	US-10-343-650A-315	Sequence 315, App
37	355.6	36.0	990	16	US-10-387-629-157	Sequence 157, App
38	355.6	36.0	1114	17	US-10-467-252-93	Sequence 93, Appl
39	355.6	36.0	1390	15	US-10-017-161-945	Sequence 945, App
40	355.6	36.0	1390	16	US-10-292-798-819	Sequence 819, App
41	351.2	35.6	1076	13	US-10-182-822A-23	Sequence 23, Appl
42	348.2	35.3	1695	15	US-10-145-586-34	Sequence 34, Appl
43	348	35.3	948	15	US-10-145-586-36	Sequence 36, Appl
44	345.2	35.0	1300	16	US-10-292-798-833	Sequence 833, App
45	340.4	34.5	948	16	US-10-387-629-189	Sequence 189, App

ALIGNMENTS

RESULT 1

US-10-081-775-1
; Sequence 1, Application US/10081775
; Publication No. US2003060409A1
; GENERAL INFORMATION:
; APPLICANT: Bristol-Myers Squibb Company
; TITLE OF INVENTION: POLYNUCLEOTIDE ENCODING A NOVEL HUMAN G-PROTEIN COUPLED
; TITLE OF INVENTION: RECEPTOR, HGPBM25, EXPRESSED HIGHLY IN IMMUNE-RELATED
; FILE REFERENCE: D0126 NP
; CURRENT APPLICATION NUMBER: US/10/081,775
; PRIOR FILING DATE: 2002-02-21
; PRIOR APPLICATION NUMBER: US 60/270,134
; PRIOR FILING DATE: 2001-02-21
; PRIOR APPLICATION NUMBER: US 60/278,952
; PRIOR FILING DATE: 2001-03-27
; NUMBER OF SEQ ID NOS: 69
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 1
; LENGTH: 1567
; TYPE: DNA
; ORGANISM: homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (537)..(1523)
US-10-081-775-1

Query Match 100.0%; Score 987; DB 15; Length 1567;
Best Local Similarity 100.0%; Pred. No. 2.6e-307;
Matches 987; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy	1	ATGTCGAGCAGCTCTTGGCCCAACATCGAATCTCTCATCATCATGATGTTGACCTTCT 60
Dd	537	ATGTCGAGCAGCTCTTGGCCCAACATCGAATCTCTCATCATCATGATGTTGACCTTCT 596
Qy	61	GTCTTCTTCTCTCTGGGCGATCCAGGCTCGAACAAATTCATTTGTGCTCTCTACTTCCCT 120
Dd	597	GTCTTCTTCTCTCTGGGCGATCCAGGCTCGAACAAATTCATTTGTGCTCTCTACTTCCCT 656

Db 841 CATATTACATCTTTTGGCCAAATGTTTATCTGCTTTTCCACCTGCTCTTAATCTGTG 900
 QY 901 GTATATGGAGTTAAGACCAACAGATCCGTAAAGAGTTGTCAAGGTGTTTCAAAGTGGG 960
 Db 901 GTATATGGAGTTAAGACCAACAGATCCGTAAAGAGTTGTCAAGGTGTTTCAAAGTGGG 960
 QY 961 CAGGGAATGGCATCAAGGCATCTGAG 987
 Db 961 CAGGGAATGGCATCAAGGCATCTGAG 987

RESULT 3

US-09-804-291-262
 ; Sequence 262, Application US/09804291
 ; Publication No. US20030088059A1
 ; GENERAL INFORMATION:
 ; APPLICANT: ZOZULIA, SERGEY
 ; TITLE OF INVENTION: HUMAN OLFACTORY RECEPTORS AND GENES ENCODING SAME
 ; FILE REFERENCE: P 0278005
 ; CURRENT APPLICATION NUMBER: US/09/804,291
 ; PRIOR FILING DATE: 2001-03-13
 ; PRIOR APPLICATION NUMBER: 60/188,914
 ; PRIOR FILING DATE: 2000-03-13
 ; PRIOR APPLICATION NUMBER: 60/192,033
 ; PRIOR FILING DATE: 2000-03-24
 ; PRIOR APPLICATION NUMBER: 60/198,474
 ; PRIOR FILING DATE: 2000-04-14
 ; PRIOR APPLICATION NUMBER: 60/199,335
 ; PRIOR FILING DATE: 2000-04-24
 ; PRIOR APPLICATION NUMBER: 60/207,702
 ; PRIOR FILING DATE: 2000-05-26
 ; PRIOR APPLICATION NUMBER: 60/213,849
 ; PRIOR FILING DATE: 2000-06-23
 ; PRIOR APPLICATION NUMBER: 60/226,534
 ; PRIOR FILING DATE: 2000-08-16
 ; PRIOR APPLICATION NUMBER: 60/230,732
 ; PRIOR FILING DATE: 2000-09-07
 ; PRIOR APPLICATION NUMBER: 60/266,862
 ; PRIOR FILING DATE: 2001-02-07
 ; NUMBER OF SEQ ID NOS: 529
 ; SOFTWARE: PatentIn Ver. 2.1
 ; SEQ ID NO 262
 ; LENGTH: 990
 ; TYPE: DNA
 ; ORGANISM: Homo sapiens
 US-09-804-291-262

Query Match 99.4%; Score 980.6; DB 10; Length 990;
 Best Local Similarity 99.6%; Pred. No. 2.3e-305;
 Matches 983; Conservative 0; Mismatches 4; Indels 0; Gaps 0;
 QY 1 ATGTCCAGCACTCTTGGCCCAACAGTGAATCTCTCATCACTGATGTTGACCCCTTCT 60
 Db 1 ATGTCCAGCACTCTTGGCCCAACAGTGAATCTCTCATCACTGATGTTGACCCCTTCT 60
 QY 61 GTCTTCTTCCCTCGGGCATCCAGGCTCTGGAACAATTTCAATTTGTGGCTCTCACTCCCT 120
 Db 61 GTCTTCTTCCCTCGGGCATCCAGGCTCTGGAACAATTTCAATTTGTGGCTCTCACTCCCT 120
 QY 121 GTGTGTGGCTTTCAGGCACAGCCCAATTTGGGCAATATAACTATTCTGGTTGTGTTGCC 180
 Db 121 GTGTGTGGCTTTCAGGCACAGCCCAATTTGGGCAATATAACTATTCTGGTTGTGTTGCC 180
 QY 181 ACTGAACCAAGCTTTCGACAGCTGTGACCTTTTCTGTGCATGCTCTCAACCATGCAC 240
 Db 181 ACTGAACCAAGCTTTCGACAGCTGTGACCTTTTCTGTGCATGCTCTCAACCATGCAC 240
 QY 241 TTGGCTGCTCTGTCTCCACAGTTCCTCAAGTACTGGTATCTTCTGTGGAGCCGGA 300
 Db 241 TTGGCTGCTCTGTCTCCACAGTTCCTCAAGTACTGGTATCTTCTGTGGAGCCGGA 300
 QY 301 CATATATCTGCTCTGCTGCTGGCACAATATGTTCTTCAATTCATGCTTCTGCATGATG 360

Db 301 CATATATCTGCTCTGCTGCTGGCACAAGATGTTCTTCAATCATGCTTCTGCATGATG 360
 QY 361 GAGTCCACCTGTCTACTGGCCCATGGCTTTGTATCGCTAGTGGCCCATCTGSCACCCACTC 420
 Db 361 GAGTCCACCTGTCTACTGGCCCATGGCTTTGTATCGTACGTGGCCCATCTGSCACCCACTC 420
 QY 421 CGGTATGCCACAATCTCTCACTGACACCATCATTTGCCACATAGGGGTGGCAGCTGTAGTG 480
 Db 421 CGGTATGCCACAATCTCTCACTGACACCATCATTTGCCACATAGGGGTGGCAGCTGTAGTG 480
 QY 481 CGAGGCTCCCTGCTCATGCTCCCATGTCCTTCTTTTATTTGGGCGTTTGAATCTTCCCAA 540
 Db 481 CGAGGCTCCCTGCTCATGCTCCCATGTCCTTCTTTTATTTGGGCGTTTGAATCTTCCCAA 540
 QY 541 AGCCATGTGATCCTTACACACAGTACTGTGAGCACATGGCTGTGTGAAGCTGCCTCTGTGA 600
 Db 541 AGCCATGTGATCCTTACACACAGTACTGTGAGCACATGGCTGTGTGAAGCTGCCTCTGTGA 600
 QY 601 GACACAGGCTTAACCGTGTGTATGGGTGACAGTGCACCTGTGTGTCATTTGGGTTGAC 660
 Db 601 GACACAGGCTTAACCGTGTGTATGGGTGACAGTGCACCTGTGTGTCATTTGGGTTGAC 660
 QY 661 TTGTTTTCATTTGCTCTCTCTATGCCCTTAATTGACAGCTGTCTTGCCTCTCATCC 720
 Db 661 TTGTTTTCATTTGCTCTCTCTATGCCCTTAATTGACAGCTGTCTTGCCTCTCATCC 720
 QY 721 CATGAAGCTCGGTCCAAGGCCCTTAGGGACCTGTGGTTCCTCATGTGTGTCTCATCTCATC 780
 Db 721 CATGAAGCTCGGTCCAAGGCCCTTAGGGACCTGTGGTTCCTCATGTGTGTCTCATCTCATC 780
 QY 781 TCTTATACACAGCCCTCTTCTCTCTTTTACACACCCGCTTGGCCATCACGTTCCAGTTC 840
 Db 781 TCTTATACACAGCCCTCTTCTCTCTTTTACACACCCGCTTGGCCATCACGTTCCAGTTC 840
 QY 841 CATATTACATTTCTTTTGGCCAAATGTTTATCTGCTTTTGGCCACCTGCTTTAATCTGTG 900
 Db 841 CATATTACATTTCTTTTGGCCAAATGTTTATCTGCTTTTGGCCACCTGCTTTAATCTGTG 900
 QY 901 GTATATGGAGTTAAGACCAACAGATCCGTAAAGAGTTGTCAAGGTGTTTCAAAGTGGG 960
 Db 901 GTATATGGAGTTAAGACCAACAGATCCGTAAAGAGTTGTCAAGGTGTTTCAAAGTGGG 960
 QY 961 CAGGGAATGGGCATCAAGGCATCTGAG 987
 Db 961 CAGGGAATGGGCATCAAGGCATCTGAG 987

RESULT 4

US-10-343-650A-353
 ; Sequence 353, Application US/10343650A
 ; Publication No. US20040067499A1
 ; GENERAL INFORMATION:
 ; APPLICANT: HAGA, TATSUYA
 ; TITLE OF INVENTION: NOVEL G-PROTEIN COUPLED RECEPTOR
 ; FILE REFERENCE: 31671-186347
 ; CURRENT APPLICATION NUMBER: US/10/343,650A
 ; CURRENT FILING DATE: 2003-07-21
 ; PRIOR APPLICATION NUMBER: JP 2000/237818
 ; PRIOR FILING DATE: 2000-08-04
 ; PRIOR APPLICATION NUMBER: JP 2001/34434
 ; PRIOR FILING DATE: 2001-02-13
 ; NUMBER OF SEQ ID NOS: 694
 ; SOFTWARE: PatentIn Ver. 2.1
 ; SEQ ID NO 353
 ; LENGTH: 990
 ; TYPE: DNA
 ; ORGANISM: Homo sapiens
 ; FEATURE:
 ; NAME/KEY: CDS
 ; LOCATION: (1)..(990)
 US-10-343-650A-353

Query Match

99.4%; Score 980.6; DB 13; Length 990;

Best Local Similarity 99.6%; Pred. No. 2.3e-305;			
Matches 983; Conservative 0; Mismatches 4; Indels 0; Gaps 0;			
QY	1	ATGCCAGCACTCTTGGCCACAAATGAAATCTCTCATCACTGATGTTGACCTTCT	60
DB	1	ATGCCAGCACTCTTGGCCACAAATGAAATCTCTCATCACTGATGTTGACCTTCT	60
QY	61	GTCTTTCTCTCTGGGATCCAGGTCTGGAACAAATTTCAATTTGGCTCTCACTCCT	120
DB	61	GTCTTTCTCTCTGGGATCCAGGTCTGGAACAAATTTCAATTTGGCTCTCACTCCT	120
QY	121	GTGTGGCTTAGGCACAGCCACAAATTTGGGCAATATAATTTCTGTTGTTGCTCC	180
DB	121	GTGTGGCTTAGGCACAGCCACAAATTTGGGCAATATAATTTCTGTTGTTGCTCC	180
QY	181	ACTGAACAGCTTTGACAAAGCTGTGTACCTTTTCTGTGCATGCTCTCAACCATGC	240
DB	181	ACTGAACAGCTTTGACAAAGCTGTGTACCTTTTCTGTGCATGCTCTCAACCATGC	240
QY	241	TTGGCTGCTCTGTCTCACAGTTCGCAAGCTACTGGCTATCTTCTGTGTGGAGCCGA	300
DB	241	TTGGCTGCTCTGTCTCACAGTTCGCAAGCTACTGGCTATCTTCTGTGTGGAGCCGA	300
QY	301	CATATATCTGCTCTGCTGCTGCTGCGCATATGTTCTTCAATCATGCTTCTGCAATG	360
DB	301	CATATATCTGCTCTGCTGCTGCTGCGCATATGTTCTTCAATCATGCTTCTGCAATG	360
QY	361	GAGTCCACTGTCTACTGCGCATGCTTTGTATGCTAGTGGCCATCTGCCACCATC	420
DB	361	GAGTCCACTGTCTACTGCGCATGCTTTGTATGCTAGTGGCCATCTGCCACCATC	420
QY	421	CGCTATGCCAATCTCTACTGACACCATCATTTGCCACATAGGGGTGGCAGCTGATG	480
DB	421	CGCTATGCCAATCTCTACTGACACCATCATTTGCCACATAGGGGTGGCAGCTGATG	480
QY	481	CGAGCTCCCTGCTCTGCTGCTGCTGCTTCTTATTTGGGCTTTGAACTTCTGCCAA	540
DB	481	CGAGCTCCCTGCTCTGCTGCTGCTGCTTCTTATTTGGGCTTTGAACTTCTGCCAA	540
QY	541	AGCCATGTGATCCTACACAGCTACTGTGAGCACATGGCTGTGGTGAAGCTGGCCTGTGA	600
DB	541	AGCCATGTGATCCTACACAGCTACTGTGAGCACATGGCTGTGGTGAAGCTGGCCTGTGA	600
QY	601	GACACAGGCTTAACGGTGTGATGGGCTGACAGCTGCACTTGTGTGATGGGTTGAC	660
DB	601	GACACAGGCTTAACGGTGTGATGGGCTGACAGCTGCACTTGTGTGATGGGTTGAC	660
QY	661	TTGTTTGTGATGCTCTCTCTATGCTTAAATTTGCAAGCTGTCTTCTGCTCTCATCC	720
DB	661	TTGTTTGTGATGCTCTCTCTATGCTTAAATTTGCAAGCTGTCTTCTGCTCTCATCC	720
QY	721	CATGAAGCTCGGTCCAGGCTTAGGACCTGTGGTTCCTCATGTCTGTGTCTCATCTC	780
DB	721	CATGAAGCTCGGTCCAGGCTTAGGACCTGTGGTTCCTCATGTCTGTGTCTCATCTC	780
QY	781	TCTTATACAGGCT	840
DB	781	TCTTATACAGGCT	840
QY	841	CATATTCATCTTTTGGCCAAATGTTTATCTCTCTTTTGGCCACCTGCTTAACTCTGTG	900
DB	841	CATATTCATCTTTTGGCCAAATGTTTATCTCTCTTTTGGCCACCTGCTTAACTCTGTG	900
QY	901	GTATATGGAGTTAAGACCAACAGATCCGTTAAAGAGTTGTCAAGGTGTTTCAAGTGGG	960
DB	901	GTATATGGAGTTAAGACCAACAGATCCGTTAAAGAGTTGTCAAGGTGTTTCAAGTGGG	960
QY	961	CAGGGAATGGCATCAAGCATCTGAG 987	
DB	961	CAGGGAATGGCATCAAGCATCTGAG 987	

US-10-025-806-35			
; Sequence 35; Application US/10025806			
; Publication No. US2003019855A1			
; GENERAL INFORMATION:			
; APPLICANT: Li, Li			
; APPLICANT: Padigaru, Muralidhara			
; APPLICANT: Ballinger, Robert			
; APPLICANT: Kekuda, Ramesh			
; APPLICANT: Colman, Steven			
; APPLICANT: Spytek, Kimberly			
; APPLICANT: Casman, Stacie			
; APPLICANT: Edinger, Valermit			
; APPLICANT: Gerlach, Valerie			
; APPLICANT: Sciore, Paul			
; APPLICANT: Smithson, Glenda			
; APPLICANT: Peyman, John			
; APPLICANT: MacDougall, John			
; APPLICANT: Stone, David			
; APPLICANT: Vernet, Corine			
; APPLICANT: Shenoy, Suresh			
; APPLICANT: Gunther, Erik			
; APPLICANT: Millet, Isabelle			
; APPLICANT: Tchernev, Velizar			
; APPLICANT: Anderson, David			
; APPLICANT: Gusev, Vladimir			
; APPLICANT: Malyankar, Uriel			
; APPLICANT: Zhong, Haibong			
; APPLICANT: Ellerman, Karen			
; APPLICANT: Wolenc, Adam			
; TITLE OF INVENTION: NOVEL PROTEINS AND NUCLEIC ACIDS ENCODING SAME			
; FILE REFERENCE: 21402-224 AB			
; CURRENT APPLICATION NUMBER: US/10/025,806			
; CURRENT FILING DATE: 2001-12-19			
; PRIOR APPLICATION NUMBER: 60/256,635			
; PRIOR FILING DATE: 2000-12-18			
; PRIOR APPLICATION NUMBER: 60/259,743			
; PRIOR FILING DATE: 2001-01-04			
; PRIOR APPLICATION NUMBER: 60/299,327			
; PRIOR FILING DATE: 2001-06-19			
; PRIOR APPLICATION NUMBER: 60/261,498			
; PRIOR FILING DATE: 2001-01-12			
; PRIOR APPLICATION NUMBER: 60/263,689			
; PRIOR FILING DATE: 2001-01-24			
; PRIOR APPLICATION NUMBER: 60/276,464			
; PRIOR FILING DATE: 2001-02-08			
; PRIOR APPLICATION NUMBER: 60/271,021			
; PRIOR FILING DATE: 2001-02-22			
; PRIOR APPLICATION NUMBER: 60/275,946			
; PRIOR FILING DATE: 2001-03-14			
; PRIOR APPLICATION NUMBER: 60/278,150			
; PRIOR FILING DATE: 2001-03-23			
; PRIOR APPLICATION NUMBER: 60/285,718			
; PRIOR FILING DATE: 2001-04-23			
; PRIOR APPLICATION NUMBER: 60/312,902			
; PRIOR FILING DATE: 2001-08-16			
; PRIOR APPLICATION NUMBER: 60/257,876			
; PRIOR FILING DATE: 2000-12-21			
; PRIOR APPLICATION NUMBER: 60/260,718			
; PRIOR FILING DATE: 2001-01-10			
; PRIOR APPLICATION NUMBER: 60/284,591			
; PRIOR FILING DATE: 2001-04-18			
; NUMBER OF SEQ ID NOS: 352			
; SOFTWARE: PatentIn Ver. 2.1			
; SEQ ID NO 35			
; LENGTH: 998			
; TYPE: DNA			
; ORGANISM: Homo sapiens			
; FEATURE:			
; NAME/KEY: CDS			
; LOCATION: (2)..(988)			
US-10-025-806-35			

Best Local Similarity 99.6%; Pred. No. 2.3e-305;			
Matches 983; Conservative 0; Mismatches 4; Indels 0; Gaps 0;			
QY	1	ATGTCAGCACTCTTGGCCGACAACTGAATCTCTCATCACACTGATGTTGACCTCTCT	60
Db	2	ATGTCAGCACTCTTGGCCGACAACTGAATCTCTCATCACACTGATGTTGACCTCTCT	61
QY	61	GTCTTCTTCTCTCTGGGATCCAGCTCTGGAACTTTCAATTTGGCTCTCATCTCCCT	120
Db	62	GTCTTCTTCTCTCTGGGATCCAGCTCTGGAACTTTCAATTTGGCTCTCATCTCCCT	121
QY	121	GTGTGTGGCTTAGGCACAGCCCAATTTGTGGCAATAAATTAATTTCTGTCTTGTGCG	180
Db	122	GTGTGTGGCTTAGGCACAGCCCAATTTGTGGCAATAAATTAATTTCTGTCTTGTGCG	181
QY	181	ACTGAACAGCTTTGCAACAGCTGTGACCTTTTCTGTGATGCTCTCAACATCGAC	240
Db	182	ACTGAACAGCTTTGCAACAGCTGTGACCTTTTCTGTGATGCTCTCAACATCGAC	241
QY	241	TTGGCTGCTCTGTCTCCACAGTTCCTCAAGCTACTGCTATCTTCTGTGTGGACCGGA	300
Db	242	TTGGCTGCTCTGTCTCCACAGTTCCTCAAGCTACTGCTATCTTCTGTGTGGACCGGA	301
QY	301	CATATATCTGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	360
Db	302	CATATATCTGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	361
QY	361	GAGTCCACTGTCTACTGGCCATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	420
Db	362	GAGTCCACTGTCTACTGGCCATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	421
QY	421	CGCTATGCAACCTCTCACTGACACCATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	480
Db	422	CGCTATGCAACCTCTCACTGACACCATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	481
QY	481	CGAGGCTCCCTGCTCATGCTCCCATGCTCCCTCTTTTATTTGGGCGTTTGAATCTTGC	540
Db	482	CGAGGCTCCCTGCTCATGCTCCCATGCTCCCTCTTTTATTTGGGCGTTTGAATCTTGC	541
QY	541	AGCATGTGATCTTACACAGCTACTGTGACACATGCTGCTGCTGCTGCTGCTGCTGCTG	600
Db	542	AGCATGTGATCTTACACAGCTACTGTGACACATGCTGCTGCTGCTGCTGCTGCTGCTG	601
QY	601	GACACAGGCTTAACCGCTGTATGGCTGACAGCTGCTGCTGCTGCTGCTGCTGCTGCTG	660
Db	602	GACACAGGCTTAACCGCTGTATGGCTGACAGCTGCTGCTGCTGCTGCTGCTGCTGCTG	661
QY	661	TTGTTTTGCAATGCTC	720
Db	662	TTGTTTTGCAATGCTC	721
QY	721	CATGAAGCTGGTCAAGGCTTACAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	780
Db	722	CATGAAGCTGGTCAAGGCTTACAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	781
QY	781	TCTTATACACAGCTC	840
Db	782	TCTTATACACAGCTC	841
QY	841	CATATTCACATCTTTTGGCCAAATGTTTATCTGCTTTTGGCCAACTGCTTCTTAACTCTG	900
Db	842	CATATTCACATCTTTTGGCCAAATGTTTATCTGCTTTTGGCCAACTGCTTCTTAACTCTG	901
QY	901	GTATATGAGTTAGACAAACAGATCGGTAAGAGTTGTCAGGCTGTTTCAAGTGGG	960
Db	902	GTATATGAGTTAGACAAACAGATCGGTAAGAGTTGTCAGGCTGTTTCAAGTGGG	961
QY	961	CAGGGAATGGGCATCAAGGCATCTGAG	987
Db	962	CAGGGAATGGGCATCAAGGCATCTGAG	988

US-10-025-806-37
; Sequence 37, Application US/10025806
; Publication No. US20030198955A1
; GENERAL INFORMATION:
; APPLICANT: Li, Li
; APPLICANT: Padigaru, Muralidhara
; APPLICANT: Ballinger, Robert
; APPLICANT: Kekuda, Ramesh
; APPLICANT: Colman, Steven
; APPLICANT: Spytek, Kimberly
; APPLICANT: Casman, Stacie
; APPLICANT: Edinger, Shlomit
; APPLICANT: Gerlach, Valerie
; APPLICANT: Sciore, Paul
; APPLICANT: Smithson, Glennda
; APPLICANT: Peyman, John
; APPLICANT: MacDougall, John
; APPLICANT: Stone, David
; APPLICANT: Vernet, Corine
; APPLICANT: Shenoy, Suresh
; APPLICANT: Gunther, Erik
; APPLICANT: Millet, Isabelle
; APPLICANT: Tchernev, Velizar
; APPLICANT: Anderson, David
; APPLICANT: Gusev, Vladimir
; APPLICANT: Malyankar, Uriel
; APPLICANT: Zhong, Haihong
; APPLICANT: Ellerman, Karen
; APPLICANT: Wolenc, Adam
; TITLE OF INVENTION: NOVEL PROTEINS AND NUCLEIC ACIDS ENCODING SAME
; FILE REFERENCE: 21402-224 AB
; CURRENT APPLICATION NUMBER: US/10/025.806
; CURRENT FILING DATE: 2001-12-19
; PRIOR APPLICATION NUMBER: 60/256,635
; PRIOR FILING DATE: 2000-12-18
; PRIOR APPLICATION NUMBER: 60/259,743
; PRIOR FILING DATE: 2001-01-04
; PRIOR APPLICATION NUMBER: 60/299,327
; PRIOR FILING DATE: 2001-06-19
; PRIOR APPLICATION NUMBER: 60/261,498
; PRIOR FILING DATE: 2001-01-12
; PRIOR APPLICATION NUMBER: 60/263,689
; PRIOR FILING DATE: 2001-01-24
; PRIOR APPLICATION NUMBER: 60/276,464
; PRIOR FILING DATE: 2001-02-08
; PRIOR APPLICATION NUMBER: 60/271,021
; PRIOR FILING DATE: 2001-02-22
; PRIOR APPLICATION NUMBER: 60/275,946
; PRIOR FILING DATE: 2001-03-14
; PRIOR APPLICATION NUMBER: 60/278,150
; PRIOR FILING DATE: 2001-03-23
; PRIOR APPLICATION NUMBER: 60/285,718
; PRIOR FILING DATE: 2001-04-23
; PRIOR APPLICATION NUMBER: 60/312,902
; PRIOR FILING DATE: 2001-08-16
; PRIOR APPLICATION NUMBER: 60/257,876
; PRIOR FILING DATE: 2000-12-21
; PRIOR APPLICATION NUMBER: 60/260,718
; PRIOR FILING DATE: 2001-01-10
; PRIOR APPLICATION NUMBER: 60/284,591
; NUMBER OF SEQ ID NOS: 352
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 37
; LENGTH: 998
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (2)..(988)
US-10-025-806-37

Best Local Similarity 99.6%; Pred. No. 2.3e-305;	
Matches	Conservative 0; Mismatches 4; Indels 0; Gaps 0;
Qy	1 ATGTCAGCACTCTTTGGCCACAACATGGAATCTCCTCATCACTGATGTTGACCCCTTCT 60
Db	2 ATGTCAGCACTCTTTGGCCACAACATGGAATCTCCTAATCACATGATGTTGACCCCTTCT 61
Qy	61 GTCTTCTTCCTCTCTGGGCATCCCAAGTCTGGAACAATTTCAATTTGGCTCTCACTCCCT 120
Db	62 GTCTTCTTCCTCTCTGGGCATCCCAAGTCTGGAACAATTTCAATTTGGCTCTCACTCCCT 121
Qy	121 GTGTGTGGCTTAGGCAACAGCCACAATTTGTGGGCAATTAACATATTCTGTGTTGTGTGGC 180
Db	122 GTGTGTGGCTTAGGCAACAGCCACAATTTGTGGGCAATTAACATATTCTGTGTTGTGTGGC 181
Qy	181 ACTGAACCAAGTCTTGACAAGCCGTGTGACCTTTTCTGTGCATGCTCTCAACCATGCA 240
Db	182 ACTGAACCAAGTCTTGACAAGCCGTGTGACCTTTTCTGTGCATGCTCTCAACCATGCA 241
Qy	241 TTGSCCTGCCTCTGTCTCCACAGTTCCCAAGTACTGCGTATCTTTCTGTGTGGAGCCGGA 300
Db	242 TTGSCCTGCCTCTGTCTCCACAGTTCCCAAGTACTGCGTATCTTTCTGTGTGGAGCCGGA 301
Qy	301 CATATATCTGCCTCTGCCTGGCCACATATGTTCTTCAATTCATGCCCTCTCGCATGATG 360
Db	302 CATATATCTGCCTCTGCCTGGCCACAGATGTTCTTCAATTCATGCCCTCTCGCATGATG 361
Qy	361 GAGTCCACTGTGCTACTGGCCATGSCCTTTTCATCGCTACGTGGCCATCTGCCACCCACTC 420
Db	362 GAGTCCACTGTGCTACTGGCCATGSCCTTTTCATCGCTACGTGGCCATCTGCCACCCACTC 421
Qy	421 CGCTATGCCACAATCCTCACTGACACCATCATATGCCCATATAGGGGTGGCAGCTGTAGTG 480
Db	422 CGCTATGCCACAATCCTCACTGACACCATCATATGCCCATATAGGGGTGGCAGCTGTAGTG 481
Qy	481 CGAGGCTCCCTGCTCATGCTCCCATGTCCCTTCTTTATTTGGCGCTTTGAACTTCTGCCAA 540
Db	482 CGAGGCTCCCTGCTCATGCTCCCATGTCCCTTCTTTATTTGGCGCTTTGAACTTCTGCCAA 541
Qy	541 AGCCATGTGATCCTTACACACAGTACTGTGAGCACATGGCTGTGGTGAAGCTGGCCTGTGGA 600
Db	542 AGCCATGTGATCCTTACACACAGTACTGTGAGCACATGGCTGTGGTGAAGCTGGCCTGTGGA 601
Qy	601 GACACCGGCCTAACCGTGTGTATGGGTGACAGCTGCATCTGTGGTCAATTTGGGGTTGAC 660
Db	602 GACACCGGCCTAACCGTGTGTATGGGTGACAGCTGCATCTGTGGTCAATTTGGGGTTGAC 661
Qy	661 TTGTTTTGCATTTGTTCTCTCTATGCCCCATATTGCACAAGCTGTGCTTGCCTCTCAATCC 720
Db	662 TTGTTTTGCATTTGTTCTCTCTATGCCCCATATTGCACAAGCTGTGCTTGCCTCTCAATCC 721
Qy	721 CATGAAGCTCGGTCCAAGGCCCTTAGGGACCTGTGGTCCCATGTCTGTGTCAATCCTCATC 780
Db	722 CATGAAGCTCGGTCCAAGGCCCTTAGGGACCTGTGGTCCCATGTCTGTGTCAATCCTCATC 781
Qy	781 TCTTTATACACGAGCCCTTCTCTCTTTTTTATACACCGCTTTGGCCATCACTGTTCCAGTC 840
Db	782 TCTTTATACACGAGCCCTTCTCTCTTTTTTATACACCGCTTTGGCCATCACTGTTCCAGTC 841
Qy	841 CATATTACATTCCTTTTGGCCAAATGTTTATCTGCTTTTGGCCACCTGTCTCTTAATCCTGTG 900
Db	842 CATATTACATTCCTTTTGGCCAAATGTTTATCTGCTTTTGGCCACCTGTCTCTTAATCCTGTG 901
Qy	901 GTATATGGAGTTAAGACCAAAACAGATCCGTAAAAGAGTTGTCAAGGCTGTTTCAAAGTGGG 960
Db	902 GTATATGGAGTTAAGACCAAAACAGATCCGTAAAAGAGTTGTCAAGGCTGTTTCAAAGTGGG 961
Qy	961 CAGGGAATGGGCATCAAGGCATCTGAG 987
Db	962 CAGGGAATGGGCATCAAGGCATCTGAG 988

RESULT 7

```

US-10-017-161-923
; Sequence 923, Application US/10017161
; Publication No. US20030143668A1
; GENERAL INFORMATION:
; APPLICANT: SUWA, MAKIKO
; APPLICANT: ASAI, KIYOSHI
; APPLICANT: AKIYAMA, YUTAKA
; APPLICANT: ABEURATANI, HIROYUKI
; TITLE OF INVENTION: NOVEL G PROTEIN-COUPLED RECEPTORS
; FILE REFERENCE: 084335/0152
; CURRENT APPLICATION NUMBER: US/10/017,161
; CURRENT FILING DATE: 2002-12-18
; PRIOR APPLICATION NUMBER: JP 2001/246789
; PRIOR FILING DATE: 2001-06-18
; NUMBER OF SEQ ID NOS: 2430
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 923
; LENGTH: 1366
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: source
; LOCATION: (1)..(1366)
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (201)..(1166)
; US-10-017-1161-923

```


Db	920	TTGGACATGTGTCCTCTCATATAGTGGCCATCTTAGCCTTCTACACAACTGTGGTCACTC	979
Qy	804	CTTTTTTACACACCGCTTTGGCACAAGTTCAGTCCATATTCACATCTTTTGGCCAA	863
Db	980	TTCAGTCAATGCACCGTAGCCGCCCATGTGCCCTCATGTCACATCTCTCTGGCAA	1039
Qy	864	TGTTTATCTGCTTTTGGCACTGCTCTTAATCTGTGGTATATGGAGTTAAGACCAACA	923
Db	1040	TTTCTATCTGCTCTTCCACCCATGGTCAATCCCAATAATCTATGCTGTCAAGACCAAGCA	1099
Qy	924	GATCCGTAAAGAGTTGTCAGGCTGTTTCAAAG	956
Db	1100	AATCCGTGAGAGCATCTTTGGAGTAATCCCAAG	1132

RESULT 12

```

US-09-864-761-3771/c
; Sequence 3771, Application US/09864761
; Patent No. US20020048763A1
; GENERAL INFORMATION:
; APPLICANT: Penn, Sharon G.
; APPLICANT: Rank, David R.
; APPLICANT: Hanzel, David K.
; APPLICANT: Chen, Wensheng
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
; FILE REFERENCE: Aemica-X-1
; CURRENT APPLICATION NUMBER: US 09/864,761
; CURRENT FILING DATE: 2001-05-23
; PRIOR APPLICATION NUMBER: US 60/180,312
; PRIOR FILING DATE: 2000-02-04
; PRIOR APPLICATION NUMBER: US 60/207,456
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: US 09/632,366
; PRIOR FILING DATE: 2000-08-03
; PRIOR APPLICATION NUMBER: GB 24263.6
; PRIOR FILING DATE: 2000-10-04
; PRIOR APPLICATION NUMBER: US 60/236,359
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: PCT/US01/00666
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00667
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00664
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00669
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00665
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00668
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00663
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00662
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00661
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00670
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: US 60/234,687
; PRIOR FILING DATE: 2000-09-21
; PRIOR APPLICATION NUMBER: US 09/608,408
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: US 09/774,203
; PRIOR FILING DATE: 2001-01-29
; NUMBER OF SEQ ID NOS: 49117
; SOFTWARE: Anomax Sequence Listing Engine vers. 1.1
; SEQ ID NO 3771
; LENGTH: 1986
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:

```


Db 61 GGCTGGAAACACCTGACATCTGGATCTCCATCCCTTTCTGCTTAGCATATACACTGGCC 120
Qy 145 ATTGTGGGCAATATACTATCTGGTGTGTGGTGGCCACTGAACACAGCTCTTGCAACAAGCCT 204
Db 121 CTGCTTGGAAACTGCACTCTCTCTCTCATATCCAGGCTGATGAGCCCTCCATGAACCC 180
Qy 205 GTGTACCTTTTCTGTGATGCTCTCAACCATGACCTGGCTGCTGTCTCCACAGTT 264
Db 181 ATGTACCTCTTTCTGGGCAATGTGGCAGCCATGACCTGTCTCTCCCTCAGCACTG 240
Qy 265 CCCAAGCTACTGGCTATCTCTGTGTGGAGCCGACATATATCTGTCTCTGCTGCTGCTG 324
Db 241 CCCAATATGCTGGCAATCTGTCTCAGGATCGGAGATAAATCTTTTGTCTGTCTG 300
Qy 325 GCACATATGTTCTTCAATCATGCTCTTGTGATGAGTCCACTGTGTCTACTGGCCATG 384
Db 301 GCCCAGATGTTCT 360
Qy 385 GCTTTGATCGTACGCTGAGGCACTGACACCATCGCTATGACCAATCCTCACTGAC 444
Db 361 GCTTTGACCGCTATGCTGCTATCTGCAAGCACTGCACTACACCAAGGCTCTGACTGG 420
Qy 445 ACCATATGCCCCATAGGGGTGGCAGCTGAGTGGAGGCTCCCTGCTCATCTCCCA 504
Db 421 TCCCTCATCAAGATGGAATGGCTGCTGTGGCCCGGCTGACACTAATGACTCCA 480
Qy 505 TGTCCCTCTTTATTTGGCGTTTGAATCTTCTGCAAGCCATGATGATGATGATGATGAT 564
Db 481 CTCCCT 540
Qy 565 TGTGACACATGCTGTGTGAAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 624
Db 541 TGTGAACATATGCTGTGTGAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 600
Qy 625 GGGCTGACAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 684
Db 601 GGCAATGCTGTGGCAATGTTTATTTGCTGTGGCTGCTGCTGCTGCTGCTGCTGCTGCT 660
Qy 685 GCCCTAATGCAAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 744
Db 661 ATCTTTATTTCTCAGGCAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 720
Qy 745 GGGACCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 804
Db 721 GGGACATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 780
Qy 805 TTTTTCACACCGCTTTGGCCATGACCTTCCAGTCCATATTCATTTCTTTTGGCCAA 864
Db 781 TCAGTCAATGACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 840
Qy 925 ATCCGTAAGAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 956
Db 901 ATCCGTAAGAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 932

RESULT 15

US-10-182-822A-25

; Sequence 25, Application US/10182822A

; Publication No. US20030211493A1

; GENERAL INFORMATION:

; APPLICANT: INCYTE GENOMICS, INC.; BAUGHN, Mariah R.

; APPLICANT: AU-YOUNG, Janice; YUE, Henry

; TITLE OF INVENTION: G-PROTEIN COUPLED RECEPTORS

; FILE REFERENCE: PI-0032 USN

; CURRENT APPLICATION NUMBER: US/10/182,822A

; CURRENT FILING DATE: 2001-02-01

; PRIOR APPLICATION NUMBER: PCT/US 01/03455

; PRIOR FILING DATE: 2001-01-02

; PRIOR APPLICATION NUMBER: US 60/180,093

; PRIOR FILING DATE: 2000-02-02
; PRIOR APPLICATION NUMBER: US 60/182,045
; NUMBER OF SEQ ID NOS: 42
; SOFTWARE: PERL Program
; SEQ ID NO 25
; LENGTH: 945
; TYPE: DNA
; ORGANISM: Homo sapiens
; NAME/KEY: misc feature
; OTHER INFORMATION: Incyte ID No: 7472036CB1
US-10-182-822A-25

Query Match 38.0%; Score 375.2; DB 13; Length 945;

Best Local Similarity 62.7%; Pred. No. 5.6e-110;

Matches 584; Conservative 0; Mismatches 348; Indels 0; Gaps 0;

Qy 25 ATGGAATCTCTCATACACTGATGTTGACCCCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 84
Db 1 ATGTCAGCCTCCAAATATACCTTAACACATCCAACTGCTTCTTGTGGTGGGATTTCCA 60
Qy 85 GGTCTGGAACAATTTCAATTTTGTGCTCTCACTCCCTGTGTGTGGCTTAGGCACAGCCACA 144
Db 61 GGCCTGGAACACCTGCACTCTGGATCTCCATCCCTTCTGCTTAGCATATACACTGGCC 120
Qy 145 ATTGTGGGCAATATACTATCTGTGTGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 204
Db 121 CTGCTTGGAAACTGCACTCTCTCTCTCATATCCAGGCTGATGAGCCCTCCATGAACCC 180
Qy 205 GTGTACCTTTTCTGTGATGCTCTCAACCATGACCTTGGCTGCTGCTCTCTCTCTCTCTCT 264
Db 181 ATGTACCTCTTTCTGGCCATGTTGGCAGCCATGACCTGCTCTCTCTCTCTCTCTCTCT 240
Qy 265 CCCAAGCTACTGGCTATCTTCTGTGTGGAGCCGACATATATGCTGCTGCTGCTGCTGCTG 324
Db 241 CCCAATATGCTGGCAATGTTTATTTGCTGTGGGATCGGAGATAAATCTTTTGTCTGTCT 300
Qy 325 GCACATATGTTCTTCAATCATGCTCTGCTGATGATGATGATGATGATGATGATGATGATG 384
Db 301 GGGCAGATGTTCTTCT 360
Qy 385 GCTTTGATCGTACGCTGAGGCACTGCAACCCACTCCGCTATGCTGCTGCTGCTGCTGCTGCT 444
Db 361 GCTTTGACCGCTATGCTGCTATCTGCAAGCCACTGCACTACACCAAGGCTCTGACTGG 420
Qy 445 ACCATATGCCCCATAGGGGTGGCAGCTGATGAGGCTCCCTGCTCATGCTCTCTCTCTCT 504
Db 421 TCCCTCATCAAGATGGAATGGAATGGAATGGAATGGAATGGAATGGAATGGAATGGAAT 480
Qy 505 TGTCCCTCTTTTATTTGGCGTTTGAATCTTCTGCAAGCCATGATGATGATGATGATGATGAT 564
Db 481 CTCCCT 540
Qy 565 TGTGACACATGCTGTGTGAAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 624
Db 541 TGTGAACATATGCTGTGTGAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 600
Qy 625 GGGCTGACAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 684
Db 601 GGCAATGCTGTGGCAATGTTTATTTGCTGTGGCTGCTGCTGCTGCTGCTGCTGCTGCT 660
Qy 685 GCCCTAATGCAAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 744
Db 661 ATCTTTATTTCTCAGGCAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 720
Qy 745 GGGACCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 804
Db 721 GGGACATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 780
Qy 805 TTTTTCACACCGCTTTGGCCATGACCTTCCAGTCCATATTCATTTCTTTTGGCCAA 864
Db 781 TCAGTCAATGACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 840

Qy	865	GTATTATCTGCTTTTGGCCACCTGCTCTTAATCCTCTGGTATATGGAGTTAAGACCAACAG	924
Db	841	TTCTATCTGCTCTTCCCAACCAATGGTCAATCCCAATAATCTATGSGTGCAAGACCAAGCAA	900
Qy	925	ATCCGTAAAGAGTGTGCAGGGTGTTCAAAG	956
Db	901	ATCCGTGAGAGCATCTTGGGAGTATCCCAAG	932

Search completed: August 27, 2004, 18:39:24
Job time : 544 secs

This Page Blank (uaprk)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 27, 2004, 15:14:54 ; Search time 4169 Seconds
(without alignments)
10261.352 Million cell updates/sec

Title: US-10-081-775-1_COPY_537_1523
Perfect score: 987
Sequence: 1 agtccagcactctggcca.....tgggcatacagcattctgag 987

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 3470272 seqs, 21671516995 residues

Total number of hits satisfying chosen parameters: 6940544

Minimum DB seq length: 0

Maximum DB Seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

GenEmbl.*

1: gb.ba.*
2: gb.htg.*
3: gb.in.*
4: gb.om.*
5: gb.ov.*
6: gb.pat.*
7: gb.ph.*
8: gb.pl.*
9: gb.pr.*
10: gb.ro.*
11: gb.st.*
12: gb.sy.*
13: gb.un.*
14: gb.vi.*
15: em.ba.*
16: em.fun.*
17: em.hum.*
18: em.in.*
19: em.mu.*
20: em.on.*
21: em.or.*
22: em.ov.*
23: em.pat.*
24: em.ph.*
25: em.pl.*
26: em.ro.*
27: em.sts.*
28: em.un.*
29: em.vi.*
30: em.htg.hum.*
31: em.htg.inv.*
32: em.htg.other.*
33: em.htg.mus.*
34: em.htg.pln.*
35: em.htg.rod.*
36: em.htg.mam.*
37: em.htg.vrt.*
38: em.sv.*
39: em.htgo.hum.*
40: em.htgo.mus.*
41: em.htgo.other.*

score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	DB	ID	Description
1	983.8	99.7	185330	2	AC124272	AC124272 Homo sapi
2	983.8	99.7	203352	2	AC111179	AC111179 Homo sapi
3	980.6	99.4	990	6	BD144452	BD144452 Novel G-p
4	980.6	99.4	998	6	AX551417	AX551417 Sequence
5	980.6	99.4	998	6	AX551419	AX551419 Sequence
6	980.6	99.4	1390	6	AX646605	AX646605 Sequence
7	980.6	99.4	1390	9	AB065534	AB065534 Homo sapi
8	980.6	99.4	172027	2	AC026090	AC026090 Homo sapi
9	980.6	99.4	176597	9	AC116156	AC116156 Homo sapi
10	956.6	96.9	963	6	AX241675	AX241675 Sequence
11	956.6	96.9	966	6	AX318261	AX318261 Sequence
12	956.6	96.9	966	6	AX448757	AX448757 Sequence
13	953.2	96.6	202761	9	AC109341	AC109341 Homo sapi
14	822.2	83.3	181386	2	AC147586	AC147586 Ocolemur
15	822.2	83.3	239335	2	AC146635	AC146635 Ocolemur
16	791	80.1	948	6	AX659373	AX659373 Sequence
17	761.4	77.1	264026	2	AC113720	AC113720 Rattus no
18	759.8	77.0	990	10	AY073012	AY073012 Mus muscu
19	759.8	77.0	191432	2	AC135109	AC135109 Mus muscu
20	747	75.7	966	10	AY317777	AY317777 Mus muscu
21	629	63.7	781	4	AY355971	AY355971 Canis fam
22	376.2	38.1	980	6	AX320273	AX320273 Sequence
23	376.2	38.1	148113	9	AC010930	AC010930 Homo sapi
24	376.2	38.1	170338	2	AC138149	AC138149 Homo sapi
25	375.2	38.0	942	6	AX241840	AX241840 Sequence
26	375.2	38.0	945	6	AX207657	AX207657 Sequence
27	375.2	38.0	945	6	AX686665	AX686665 Sequence
28	375.2	38.0	945	6	BD144356	BD144356 Novel G-p
29	374.6	38.0	1345	6	AX646565	AX646565 Sequence
30	374.6	38.0	1345	9	AB065791	AB065791 Homo sapi
31	374.6	38.0	141485	9	AC009758	AC009758 Homo sapi
32	373.6	37.9	945	6	AX448711	AX448711 Sequence
33	368.4	37.3	1013	6	AX551407	AX551407 Sequence
34	368.4	37.3	1360	6	AX646629	AX646629 Sequence
35	368.4	37.3	1360	9	AB065820	AB065820 Homo sapi
36	368.4	37.3	165682	9	AC111177	AC111177 Homo sapi
37	368.4	37.3	165729	2	AC009642	AC009642 Homo sapi
38	367.2	37.2	982	6	AX320269	AX320269 Sequence
39	367.2	37.2	235030	2	AC107331	AC107331 Rattus no
40	365	37.0	1345	6	AX646563	AX646563 Sequence
41	365	37.0	1345	9	AB065790	AB065790 Homo sapi
42	364	36.9	942	6	AX241838	AX241838 Sequence
43	364	36.9	945	6	AX448713	AX448713 Sequence
44	364	36.9	1151	6	AX320267	AX320267 Sequence
45	362.4	36.7	945	6	BD144355	BD144355 Novel G-p

ALIGNMENTS

RESULT 1
AC124272
LOCUS AC124272 185330 bp DNA linear HTG 14-SEP-2002
DEFINITION Homo sapiens chromosome 11 clone RP11-607K3 map 11, WORKING DRAFT
SEQUENCE, 4 unordered pieces.
ACCESSION AC124272
VERSION AC124272.2 GI:22857733
KEYWORDS HTG; HTGS_PHASE1; HTGS_DRAFT; HTGS_FULLTOP.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 185330)
AUTHORS Birren, B., Nusbaum, C. and Lander, E.
TITLE Homo sapiens chromosome 11, clone RP11-607K3

Pred. No. is the number of results predicted by chance to have a


```

Db      3779  CGAGGCTCCCTCATGCTCCCATGCTCCCTCTTTATTTGGCGTTTGAATCTTGCCTCA 3838
QY      541  ACCCATGATCTTACACAGTACTGTGTAGACATGGCTGTGGTGAAGTGGCCCTGTGGA 600
Db      3839  AGCCATGTGATCTTACACAGTACTGTGTAGACATGGCTGTGGTGAAGTGGCCCTGTGGA 3898
QY      601  GACACAGGCTTAAACCGTGTGTATGGCTGTAGAGCTGGACATGCTTGGTCAATTTGGGCTTGC 560
Db      3899  GACACAGGCTTAAACCGTGTGTATGGCTGTAGAGCTGGACATGCTTGGTCAATTTGGGCTTGC 3958
QY      661  TTGTTTTCATGTGCTCTCTCTATGCTTAAATGTCACAAAGCTGTCTTGGCTCTCATCC 720
Db      3959  TTGTTTTCATGTGCTCTCTCTATGCTTAAATGTCACAAAGCTGTCTTGGCTCTCATCC 4018
QY      721  CATGAAGCTCGTCAAGGCCCTTAGGACCTGTGTGTTCCCATGTGTGTGTCATCCCTCATC 780
Db      4019  CATGAAGCTCGTCAAGGCCCTTAGGACCTGTGTGTTCCCATGTGTGTGTCATCCCTCATC 4078
QY      781  TCTTATACACAGGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 840
Db      4079  TCTTATACACAGGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 4138
QY      841  CATATTCATCTTTTGGCCCAATGTTATCTGCTTTTGGCCACCTGCTCTTAATCTCTGTG 900
Db      4139  CATATTCATCTTTTGGCCCAATGTTATCTGCTTTTGGCCACCTGCTCTTAATCTCTGTG 4198
QY      901  GTATATGAGTTAGACAAACAGATCCGTAAGAGTTGTCAAGGTGTTCAAGTGGG 960
Db      4199  GTATATGAGTTAGACAAACAGATCCGTAAGAGTTGTCAAGGTGTTCAAGTGGG 4258
QY      961  CAGGAATGGGCATCAAGGCATCTGAG 987
Db      4259  CAGGAATGGGCATCAAGGCATCTGAG 4285

RESULT 2
AC111179 LOCUS
DEFINITION Homo sapiens chromosome 11 clone RP11-589G14 map 11, WORKING DRAFT
AC111179
AC111179.2 GI:21699469
VERSION HTG; HTGS PHASE2; HTGS_DRAFT; HTGS_FULLTOP.
KEYWORDS Homo sapiens (human)
SOURCE Homo sapiens
ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 203352)
AUTHORS Birren,B., Nussbaum,C. and Lander,E.
TITLE Homo sapiens chromosome 11, clone RP11-589G14
JOURNAL Unpublished
AUTHORS 2 (bases 1 to 203352)
Birren,B., Linton,L., Nussbaum,C., Lander,E., Ali,A., Allen,N.,
Anderson,S., Barna,N., Bastien,V., Bloom,T., Boguslavsky,L., Boukhalter,B.,
Brown,A., Camarata,J., Campiano,A., Chang,J., Chazaro,B.,
Coepele,Y., Colangelo,M., Collins,S., Collymore,A., Cook,A.,
Cooke,P., DeRellano,K., Dewar,K., Diaz,J.S., Dodge,S., Faro,S.,
Ferreira,P., Fitzhugh,W., Gage,D., Galagan,J., Gardyna,S.,
Gardyna,S., Gordon,L., Hulme,W., Iliev,I., Johnson,R., Jones,C., Kamat,A.,
Karatas,A., Kellis,C., Landers,T., Levine,R., Lindblad-Toh,K.,
Macdonald,P., Major,J., Marquis,N., Matthews,C., McCarthy,M.,
McEwan,P., McKernan,K., Meldrum,J., Meneus,L., Mihova,T.,
Mlenga,V., Murphy,T., Naylor,J., Nguyen,C., Nicol,R., Norbu,C., Rogov,P.,
Roman,J., Roy,A., Schauer,S., Schuback,R., Seaman,S., Severy,P.,
Smith,C., Spencer,B., Stange-Thomann,N., Stojanovic,N., Talamas,J.,
Tesfaye,S., Theodore,J., Topham,K., Travers,M., Vassiliev,H.,
Viel,R., Vo,A., Wilson,B., Wu,X., Wyman,D., Young,G., Zainoun,J.,
Zembek,L., Zimmer,A. and Zody,M.

```

```

Viel,R., Vo,A., Wilson,B., Wu,X., Wyman,D., Ye,W.J., Young,G.,
Zainoun,J., Zembek,L., Zimmer,A. and Zody,M.
Direct Submission
Submitted (18-FEB-2002) Whitehead Institute/MIT Center for Genome
Research, 320 Charles Street, Cambridge, MA 02141, USA
3 (bases 1 to 203352)
Birren,B., Nussbaum,C., Lander,E., Ali,A., Allen,N., Anderson,S.,
Barna,N., Bastien,V., Bloom,T., Boguslavsky,L., Boukhalter,B.,
Camarata,J., Chang,J., Chazaro,B., Coepel,Y., Collymore,A.,
Cook,A., Cooke,P., DeRellano,K., Dewar,K., Diaz,J.S., Dodge,S.,
Faro,S., Ferreira,P., FitzGerald,M., Gage,D., Galagan,J.,
Gardyna,S., Gordon,L., Hulme,W., Iliev,I., Johnson,R., Jones,C., Kamat,A.,
Karatas,A., Kellis,C., Landers,T., Levine,R., Lindblad-Toh,K.,
Macdonald,P., Major,J., Marquis,N., Matthews,C., McCarthy,M.,
McEwan,P., McKernan,K., Meldrum,J., Meneus,L., Mihova,T.,
Mlenga,V., Murphy,T., Naylor,J., Nguyen,C., Nicol,R., Norbu,C., Rogov,P.,
Roman,J., Roy,A., Schauer,S., Schuback,R., Seaman,S., Severy,P.,
Smith,C., Spencer,B., Stange-Thomann,N., Stojanovic,N., Talamas,J.,
Tesfaye,S., Theodore,J., Topham,K., Travers,M., Vassiliev,H.,
Viel,R., Vo,A., Wilson,B., Wu,X., Wyman,D., Young,G., Zainoun,J.,
Zembek,L., Zimmer,A. and Zody,M.
Direct Submission
Submitted (08-AUG-2002) Whitehead Institute/MIT Center for Genome
Research, 320 Charles Street, Cambridge, MA 02141, USA
On Jul 5, 2002 this sequence version replaced gi:18699944.
All repeats were identified using RepeatMasker:
Smit, A.F.A. & Green, P. (1996-1997)
http://ftp.genome.washington.edu/RM/RepeatMasker.html
----- Genome Center
Center: Whitehead Institute/ MIT Center for Genome Research
Web site: http://www-seq.wi.mit.edu
Contact: sequence.submissions@genome.wi.mit.edu
----- Project Information
Center project name: L25440
Center clone name: 589.G.14
----- Summary Statistics
Sequencing vector: plasmid; n/a; 100% of reads
Chemistry: Dye-terminator Big Dye; 100% of reads
Assembly program: Phrap; version 0.960731
Consensus quality: 201894 bases at least Q40
Consensus quality: 202347 bases at least Q30
Consensus quality: 202672 bases at least Q20
Insert size: 200000; agarose-fp
Insert size: 203052; sum-of-contigs
Quality coverage: 14.6 in Q20 bases; agarose-fp
Quality coverage: 14.4 in Q20 bases; sum-of-contigs
-----
* NOTE: This is a 'working draft' sequence. It currently
* consists of 4 contigs. Gaps between the contigs
* are represented as runs of N. The order of the pieces
* is believed to be correct as given, however the sizes
* of the gaps between them are based on estimates that have
* been provided by the submitter.
* This sequence will be replaced
* by the finished sequence as soon as it is available and
* the accession number will be preserved
1 70426: contig of 70426 bp in length
70527 70526: gap of 100 bp
70537 78730: contig of 8204 bp in length
78731 78830: gap of 100 bp
78831 137440: contig of 58610 bp in length
137441 137540: gap of 100 bp
137541 203352: contig of 65812 bp in length.
-----
Location/Qualifiers
1. 203352
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/chromosome="11"

```

```
/map="11"
/clone="RP11-589G14"
/clone_lib="RPC1-11 Human Male BAC"
1..70426
/note="assembly_fragment
clone_end:SP6
vector_side:left"
70527..78730
/note="assembly_fragment"
78831..137440
/note="assembly_fragment"
137541..203352
/note="assembly_fragment
clone_end:77
vector_side:right"

misc_feature
misc_feature
misc_feature
misc_feature

ORIGIN
Query Match          99.7%; Score 983.8; DB 2; Length 203352;
Best Local Similarity 99.8%; Pred. No. 2.7e-287;
Matches 985; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 ATGTCCAGCACTTTGGCCACAACATGGAATCTCTCATCACTGATGTTGACCTTCT 60
DB 169277 ATGTCCAGCACTTTGGCCACAACATGGAATCTCTCATCACTGATGTTGACCTTCT 169336

QY 61 GTCTTCTTCTCTGGGATCCAGGCTGGNACAAATTTTCATTTGGGCTCTCACTCCCT 120
DB 169337 GTCTTCTTCTCTGGGATCCAGGCTGGNACAAATTTTCATTTGGGCTCTCACTCCCT 169396

QY 121 GTGTGTGCTTAGGCAGACGACCAATTTGGGCAATATAACTATTTCTGGTGTGTGTC 180
DB 169397 GTGTGTGCTTAGGCAGACGACCAATTTGGGCAATATAACTATTTCTGGTGTGTGTC 169456

QY 181 ACTGAACAGCTTTCGCAAGCCTGTGACCTTTTCTGTGCATGCTCTCAACCATCGAC 240
DB 169457 ACTGAACAGCTTTCGCAAGCCTGTGACCTTTTCTGTGCATGCTCTCAACCATCGAC 169516

QY 241 TTGGCTGCTCTGTCTCCACAGTCCCAAGCTACTGGCTATCTTCTGGTGGGACCGGA 300
DB 169517 TTGGCTGCTCTGTCTCCACAGTCCCAAGCTACTGGCTATCTTCTGGTGGGACCGGA 169576

QY 301 CATATATCTGCTCTGCTGCTGGCCACATATGTTCTTCAATCATGCTCTTGCATGATG 360
DB 169577 CATATATCTGCTCTGCTGCTGGCCACATATGTTCTTCAATCATGCTCTTGCATGATG 169636

QY 361 GAGTCCACTGTGCTACTGCGCCATGCGCTTTGATCGCTAGTGGCCATCTGCCACCACTC 420
DB 169637 GAGTCCACTGTGCTACTGCGCCATGCGCTTTGATCGCTAGTGGCCATCTGCCACCACTC 169696

QY 421 CGCTATGCCAATCCTCACTGACACCATCATTTGCCACATAGGCGTGGCAGCTGTAGTG 480
DB 169697 CGCTATGCCAATCCTCACTGACACCATCATTTGCCACATAGGCGTGGCAGCTGTAGTG 169756

QY 481 CGAGGCTCCCTGCTCATGCTCCCATGTCCTCTTTTATTTGGCGTTTGAACCTTCTGCCAA 540
DB 169757 CGAGGCTCCCTGCTCATGCTCCCATGTCCTCTTTTATTTGGCGTTTGAACCTTCTGCCAA 169816

QY 541 AGCCATGTGATCCTACACAGTACTGTGAGACATGCTGGTGTGGTGAAGCTGCGCTGTGA 600
DB 169817 AGCCATGTGATCCTACACAGTACTGTGAGACATGCTGGTGTGGTGAAGCTGCGCTGTGA 169876

QY 601 GACACAGGCTTAACCGTGTGATGGGCTGACAGCTGCACTGTGTGTCAATTTGGGTTGAC 660
DB 169877 GACACAGGCTTAACCGTGTGATGGGCTGACAGCTGCACTGTGTGTCAATTTGGGTTGAC 169936

QY 661 TTGTTTTGCAATGGTCTCTCTATGCTCCCTAAATTTGACAGCTGTCTCTGCGCTCTCATCC 720
DB 169937 TTGTTTTGCAATGGTCTCTCTATGCTCCCTAAATTTGACAGCTGTCTCTGCGCTCTCATCC 169996

QY 721 CATGAAGCTCGGTCCAAGGCCCTAGGACCTGTGGTCCCATGTCTGTGTCTATCTCATC 780
DB 169997 CATGAAGCTCGGTCCAAGGCCCTAGGACCTGTGGTCCCATGTCTGTGTCTATCTCATC 170056

QY 781 TCTTATACACAGCCCTCTTCTCTCTTTTATACACACCGCTTTGGCCATCAGTTCAGTC 840
DB 170057 TCTTATACACAGCCCTCTTCTCTCTTTTATACACACCGCTTTGGCCATCAGTTCAGTC 170116

QY 841 CATATTACATTTCTTTTGGCCAAATGTTTATCTGTTTGGCCACCTGCTCTTAATCTGTG 900
DB 170117 CATATTACATTTCTTTTGGCCAAATGTTTATCTGTTTGGCCACCTGCTCTTAATCTGTG 170176

QY 901 GTATATGAGTTAAGACCAACACATCGTAAAGAGTTGTTCAGGGTGTTCATAAGTGGG 960
DB 170177 GTATATGAGTTAAGACCAACACATCGTAAAGAGTTGTTCAGGGTGTTCATAAGTGGG 170236

QY 961 CAGGAATGGGCATCAGGCATCTGAG 987
DB 170237 CAGGAATGGGCATCAGGCATCTGAG 170263

RESULT 3
BD144452
LOCUS BD144452 990 bp DNA linear PAT 17-JAN-2003
DEFINITION Novel G-protein coupled receptors.
ACCESSION BD144452
VERSION BD144452.1 GI:27850210
KEYWORDS JP 2002112793-A/177.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
1 (Bases 1 to 990)
AUTHORS Haga, T., Takeda, S. and Miyake, N.
TITLE Novel G-protein coupled receptors
JOURNAL Patent: JP 2002112793-A 177 16-APR-2002;
JAPAN SCIENCE AND TECHNOLOGY CORP
COMMENT
OS Homo sapiens (human)
PN JP 2002112793-A/177
PF 16-APR-2002
PI 09-FEB-2001 JP 2001034434
PI TATSUYA HAGA, SHIGEKI TAKEDA, NARIKI MIYAKE
PC
C12N15/09, A01K67/027, A61K38/00, A61K39/395, A61K39/395, A61K45/00, PC
A61K48/00,
PC A61P43/00, C07K14/705, C07K16/28, C07K19/00, C12N1/15, C12N1/19, PC
C12N1/21,
PC C12N5/10, C12Q1/02, C12Q1/68, G01N33/15, G01N33/50, G01N33/53, PC
G01N33/566//
PC C12P21/08, C12N15/00, A61K37/02, C12N5/00
CC Novel G-protein coupled receptors
FH Key Location/Qualifiers
FT CDS
(1)..(990).
Location/Qualifiers
1..990
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"

FEATURES
ORIGIN
Query Match 99.4%; Score 980.6; DB 6; Length 990;
Best Local Similarity 99.6%; Pred. No. 1.3e-286;
Matches 983; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 ATGTCCAGCACTTTGGCCACAACATGGAATCTCTCATCACTGATGTTGACCTTCT 60
DB 1 ATGTCCAGCACTTTGGCCACAACATGGAATCTCTCATCACTGATGTTGACCTTCT 60

QY 61 GTCTTCTTCTCTGGGATCCAGGCTGGNACAAATTTTCATTTGGGCTCTCACTCCCT 120
DB 61 GTCTTCTTCTCTGGGATCCAGGCTGGNACAAATTTTCATTTGGGCTCTCACTCCCT 120

QY 121 GTGTGTGCTTAGGCAGACGACCAATTTGGGCAATATAACTATTTCTGGTGTGTGTC 180
DB 121 GTGTGTGCTTAGGCAGACGACCAATTTGGGCAATATAACTATTTCTGGTGTGTGTC 180

QY 181 ACTGAACAGCTTTCGCAAGCCTGTGACCTTTTCTGTGCATGCTCTCAACCATCGAC 240
```


QY 721 CATGAAGCTCGTCCAGGCGCTAGGACCTGCTGCTCCCATGCTGCTGCTCATCTCTATC 780
 Db 722 CATGAAGCTCGTCCAGGCGCTAGGACCTGCTGCTCCCATGCTGCTGCTCATCTCTATC 781
 QY 781 TCTTATACACAGGCGCTCTCTCTCTTTTATACACACCGCTTTGGCCATCAGCTTCCAGTC 840
 Db 782 TCTTATACACAGGCGCTCTCTCTCTTTTATACACACCGCTTTGGCCATCAGCTTCCAGTC 841
 QY 841 CATATTCACATCTTTTGGCCAAATGTTATCTGCTTTTGGCCACCTGCTCTTAATCCTGTG 900
 Db 842 CATATTCACATCTTTTGGCCAAATGTTATCTGCTTTTGGCCACCTGCTCTTAATCCTGTG 901
 QY 901 GTATATGAGTTAGACCAACAGATCCGTAAAGAGTTGTCAGGGTGTTCAAAGTGGG 960
 Db 902 GTATATGAGTTAGACCAACAGATCCGTAAAGAGTTGTCAGGGTGTTCAAAGTGGG 961
 QY 961 CAGGGAATGGGCATCAAGGCATCTGAG 987
 Db 962 CAGGGAATGGGCATCAAGGCATCTGAG 988

RESULT 5
 AX551419 998 bp DNA linear PAT 26-NOV-2002
 LOCUS Sequence 37 from Patent WO0250276.
 DEFINITION AX551419
 ACCESSION AX551419.1 GI:25814219
 VERSION
 KEYWORDS
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 1
 Li, L., Padigaru, M., Ballinger, R.A., Kekuda, R., Colman, S.D.,
 Scire, P., Smithson, G., Peyman, J.A., Macdougall, J.R., Stone, D.,
 Vernet, C.A., Shenoy, S., Gunther, E., Millett, I., Tchernev, V.T.,
 Anderson, D., Gusev, V., Malyankar, U.M., Zhong, H., Ellerman, K.E. and
 Wolenc, A.
 Novel proteins and nucleic acids encoding same
 Patent: WO 0250276-A 37 27-JUN-2002;
 Curagen Corporation (US)
 FEATURES
 Location/Qualifiers
 1..998
 /organism="Homo sapiens"
 /mol_type="unassigned DNA"
 /db_xref="taxon:9606"
 2..991
 /note="unnamed protein product"
 /codon_start=1
 /protein_id="CAD57921.1"
 /db_xref="GI:25814220"
 /translation="MSSTLGHNESPNDVDPVFFLLIGLPGLEQFHLWLSLPVCGL
 GATATVNIIVWATEPVLHPVLFCLMSTLDLAASVSTVPKLAIFWCGAGHI
 SASCLAOFFTHAFOMESTVLLAMAPDRVAICHPLRYATILDTIIAHIGVAIV
 RGLSLMLCPFLIGLNFQSHVILHLYCEHWAVKLAGDTRPNRVGLTRALLVIG
 VDFICGLYSALSAQVRLRSLSHRSKALCTGSHVGVILLISYTPALFSFTHRFH
 HVPVHILLANVLLLLPALPNVYGVYTKQIRKRVVRVFGQGMGINKASE"

TITLE
 JOURNAL
 Curagen Corporation (US)

FEATURES
 source

CDS

ORIGIN

Query Match 99.4%; Score 980.6; DB 6; Length 998;
 Best Local Similarity 99.6%; Pred. No. 1.3e-286;
 Matches 983; Conservative 0; Mismatches 4; Indels 0; Gaps 0;
 QY 1 ATGTCCAGCACTCTTGGCCCAACATGGAATCTCCATCACAATGATGTTGACCTTCT 60
 Db 2 ATGTCCAGCACTCTTGGCCCAACATGGAATCTCCATCACAATGATGTTGACCTTCT 61
 QY 61 GTCCTTCTCTCTGGGCATCCAGGCTCTGGAACATTTTCATTTGGGCTCTCACCTCC 120
 Db 62 GTCCTTCTCTCTGGGCATCCAGGCTCTGGAACATTTTCATTTGGGCTCTCACCTCC 121

QY 121 GTGTGTGCTTAGGCACAGCCACAATTTGTGGCAATATAAATAATTTCTGGTTGTTGTC 180
 Db 122 GTGTGTGCTTAGGCACAGCCACAATTTGTGGCAATATAAATAATTTCTGGTTGTTGTC 181
 QY 181 ACTGAACAGTCTTTGGCAAGCCTGTGTACCTTTTCTGTGCATGCTCTCAACCATCCAC 240
 Db 182 ACTGAACAGTCTTTGGCAAGCCTGTGTACCTTTTCTGTGCATGCTCTCAACCATCCAC 241
 QY 241 TTGGTGCTCTGTCTCCACAGTTCCCAAGCTACTGGCTATCTTCTGGTGTGGAGCCGGA 300
 Db 242 TTGGTGCTCTGTCTCCACAGTTCCCAAGCTACTGGCTATCTTCTGGTGTGGAGCCGGA 301
 QY 301 CATATATCTGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 360
 Db 302 CATATATCTGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 361
 QY 361 GAGTCCACTGTGCTACTGGCCATGGCCCTTTGATCGCTACGTGGCCATCTGCCACCCTC 420
 Db 362 GAGTCCACTGTGCTACTGGCCATGGCCCTTTGATCGCTACGTGGCCATCTGCCACCCTC 421
 QY 421 CGCTATGCCACAATCTCTCACTGACACCATCATTTGCCACATAGGCGGTGGCAGCTAG 480
 Db 422 CGCTATGCCACAATCTCTCACTGACACCATCATTTGCCACATAGGCGGTGGCAGCTAG 481
 QY 481 CGAGGCTCCCTGCTCATGCTCCCATGCTCCCTTTTATTTGGGCGTTTGAATCTTCTGCC 540
 Db 482 CGAGGCTCCCTGCTCATGCTCCCATGCTCCCTTTTATTTGGGCGTTTGAATCTTCTGCC 541
 QY 541 AGCCATGTGATCTTACACAGCTACTGTGAGCAGACATGGCTGTGTTGAAGTGGCCTGTG 600
 Db 542 AGCCATGTGATCTTACACAGCTACTGTGAGCAGACATGGCTGTGTTGAAGTGGCCTGTG 601
 QY 601 GACACAGGCGCTTAACCGGTGTATGGGCTGACAGCTGCACTGTTGGTCAATTTGGGTTGAC 660
 Db 602 GACACAGGCGCTTAACCGGTGTATGGGCTGACAGCTGCACTGTTGGTCAATTTGGGTTGAC 661
 QY 661 TTGTTTTGCATTTGCTCTCTCTATGCTCCCTAAATTTGACAGCTGCTCTGCGCTCTCATCC 720
 Db 662 TTGTTTTGCATTTGCTCTCTCTATGCTCCCTAAATTTGACAGCTGCTCTCTGCGCTCTCATCC 721
 QY 721 CATGAAGCTCGTCCAGGCGCTAGGACCTGCTGCTGCCATGCTGTTGCTCATCTCTCATC 780
 Db 722 CATGAAGCTCGTCCAGGCGCTAGGACCTGCTGCTGCCATGCTGTTGCTCATCTCTCATC 781
 QY 781 TCTTATACACAGCGCTCTTCTCTCTTTTACACACCGCTTTGGCCATCAGCTTCCAGTC 840
 Db 782 TCTTATACACAGCGCTCTTCTCTCTTTTACACACCGCTTTGGCCATCAGCTTCCAGTC 841
 QY 841 CATATTCACATCTTTTGGCCAAATGTTATCTGCTTTTGGCCACCTGCTCTTAATCCTGTG 900
 Db 842 CATATTCACATCTTTTGGCCAAATGTTATCTGCTTTTGGCCACCTGCTCTTAATCCTGTG 901
 QY 901 GTATATGAGTTAGACCAACAGATCCGTAAAGAGTTGTCAGGGTGTTCAAAGTGGG 960
 Db 902 GTATATGAGTTAGACCAACAGATCCGTAAAGAGTTGTCAGGGTGTTCAAAGTGGG 961
 QY 961 CAGGGAATGGGCATCAAGGCATCTGAG 987
 Db 962 CAGGGAATGGGCATCAAGGCATCTGAG 988

RESULT 6

AX646605

LOCUS

Sequence 797 from Patent EPI270724.

DEFINITION

AX646605

ACCESSION

AX646605.1

GI:28798988

KEYWORDS

SOURCE

Homo sapiens (human)

ORGANISM

REFERENCE

1

AX646605

Sequence 797 from Patent EPI270724.

DEFINITION

AX646605

ACCESSION

AX646605.1

GI:28798988

KEYWORDS

SOURCE

Homo sapiens (human)

ORGANISM

REFERENCE

1

AX646605

Sequence 797 from Patent EPI270724.

DEFINITION

AX646605

ACCESSION

AX646605.1

GI:28798988

KEYWORDS

SOURCE

Homo sapiens (human)

ORGANISM

REFERENCE

1

AUTHORS Suwa, M., Asai, K., Akiyama, Y. and Aburatani, H.
TITLE Guanosine triphosphate-binding protein coupled receptors
JOURNAL Patent: EP 1270724-A 797 02-JAN-2003;
National Institute of Advanced Industrial Science and Technology
(JP) ; Center for Advanced Science and Technology Incubation, Ltd.
(JP)

FEATURES	source
Location/Qualifiers	
1..11390	
/organism="Homo sapiens"	
/mol_type="unassigned DNA"	
/db_xref="taxon:9606"	
201..1190	
/note="unnamed protein product"	
/codon_start=1	
/protein_id="CAD69429.1"	
/db_xref="GI:28798989"	
/db_xref="REMBL:CAD69429"	
/translation="MSSTLGHNMESPNHTVDVSVFLLGIPGLEPHLMWLSIPVCGLI GTATIVGNIIILVVVATEPVLHKPVIYDFLCMLSTIDLAASVSTVPKLLAIFWCGAGHI RSAGLACMPFIIHACGMESTVILAMAFDHYAICHELRYATILTTITTAHIGVAAVY RGSSLLMPCPLFIRLNCQSHVILTYCEHMAVFWKLACGDTPRNRYVGLTAALLVIG VDFICIGLSVALSAQVLRILSHSEARSKALGTGSHVCVILISYTPALRSFFTHRFGRG HVPVHIILLANYVLLPPALNPVYGVKTKOTRKRVRVYFVESCOCMCCTKASE"	

ORIGIN

Query Match	99.4%;	Score 980.6;	DB 6;	Length 1390;
Best Local Similarity	99.6%;	Pred. No. 1.4e-286;		
Matches 983; Conservative	0;	Mismatches 4;	Indels 0;	Gaps 0;
QY	1	ATGTCAGACACTCTTGGCCACAACATGGAATCTCCCTCATCACACTGATGTGGACCTTCT	60	
Db	201	ATGTCAGACACTCTTGGCCACAACATGGAATCTCCCTCATCACACTGATGTGGACCTTCT	260	
QY	61	GTCTTCTTCTCTCTGGGCATCCCAAGTCTGTGGAACAATTCATTTGGTGGCTCTCACTCCCT	120	
Db	261	GTCTTCTTCTCTCTGGGCATCCCAAGTCTGTGGAACAATTCATTTGGTGGCTCTCACTCCCT	320	
QY	121	GTGTGTGGCTTAGGCACAGCCACAATTTGTGGCAATATAACTATTCTGGTTGTGTGGC	180	
Db	321	GTGTGTGGCTTAGGCACAGCCACAATTTGTGGCAATATAACTATTCTGGTTGTGTGGC	380	
QY	181	ACTGAACCAAGTCTTGACAAGCCTGTGTACCTTTTCTGTGCATGCTCTCAACCATCGAC	240	
Db	381	ACTGAACCAAGTCTTGACAAGCCTGTGTACCTTTTCTGTGCATGCTCTCAACCATCGAC	440	
QY	241	TTGGGTCGCTCTGTCTCCACAGTTCCTCAAGCTACTGGCTATCTTCTGGGTGGAGCCGA	300	
Db	441	TTGGGTCGCTCTGTCTCCACAGTTCCTCAAGCTACTGGCTATCTTCTGGGTGGAGCCGA	500	
QY	301	CATATATCTGCCTCTGCCTGCCTGGCACATATGTTCTTCATTATGCTCCCTCTGCGATGATG	360	
Db	501	CATATATCTGCCTCTGCCTGCCTGGCACAGATGTTCTTCATTATGCTCCCTCTGCGATGATG	560	
QY	361	GAGTCCACGTGCTACTGCGCATGGCCTTTGATCGCTACGTGGCCATCTGCCACCATTC	420	
Db	561	GAGTCCACGTGCTACTGCGCATGGCCTTTGATCGCTACGTGGCCATCTGCCACCATTC	620	
QY	421	CGCTATGGCCAAATCCTCACTGACACCATCATTTGGCCACATAGGGGTGGCAGCTGTAGT	480	
Db	621	CGCTATGGCCAAATCCTCACTGACACCATCATTTGGCCACATAGGGGTGGCAGCTGTAGT	680	
QY	481	CGAGGCTCCCTGCTCATGCTCCCATGTCCCTTTCTTTATTGGGCGGTTTGAACCTCTGCGCAA	540	
Db	681	CGAGGCTCCCTGCTCATGCTCCCATGTCCCTTTCTTTATTGGGCGGTTTGAACCTCTGCGCAA	740	
QY	541	AGCCATGTGATCCTACACAGCTACTGTGAGCACATGGCTGTGGTGAAGCTGSCCTGTGGA	600	
Db	741	AGCCATGTGATCCTACACAGCTACTGTGAGCACATGGCTGTGGTGAAGCTGSCCTGTGGA	800	
QY	601	GACACAGGCGCTAACCGTGTGTATGGGTGACAGCTGACACTGTTGGTCAATGGGGTTGAC	660	
Db	801	GACACAGGCGCTAACCGTGTGTATGGGTGACAGCTGACACTGTTGGTCAATGGGGTTGAC	860	

Qy	661	TTGTTTGGCAATGGTCTCTCTATGCGCCTAATTGCAACAGTGTCTCTTGCCTCTCATCC	720
Db	861	TTGTTTGGCAATGGTCTCTCTATGCGCCTAAGTGCAACAGTGTCTCTTGCCTCTCATCC	920
Qy	721	CATGAAGCTCGGTCCAAAGGCCCTAGGACCTGTGTTCCCATGTCTGTGTCACTCTCATC	780
Db	921	CATGAAGCTCGGTCCAAAGGCCCTAGGACCTGTGTTCCCATGTCTGTGTCACTCTCATC	980
Qy	781	TCATTATACACGACCCCTCTCTCCTTTTATACACGGCTTTGGCCATCAGGTTCCAGTC	840
Db	981	TCATTATACACGACCCCTCTCTCCTTTTATACACGGCTTTGGCCATCAGGTTCCAGTC	1040
Qy	841	CATATTACAAATCTTTTGGCCAAATGTTATCTCTTTTGCACCTGCTCTTAATCCTGTG	900
Db	1041	CATATTACAAATCTTTTGGCCAAATGTTATCTCTTTTGCACCTGCTCTTAATCCTGTG	1100
Qy	901	GTATATGGAGTTAAGACCAACAGATCCGTAAAGAGTTGTCTCAGGGTGTTTCAAAGTGGG	960
Db	1101	GTATATGGAGTTAAGACCAACAGATCCGTAAAGAGTTGTCTCAGGGTGTTTCAAAGTGGG	1160
Qy	961	CAGGGAATGGGCATCAAGGCATCTGAG	987
Db	1161	CAGGGAATGGGCATCAAGGCATCTGAG	1187

RESULT 7	
LOCUS	AB065534
DEFINITION	Homo sapiens gene for seven transmembrane helix receptor, complete cds, isolate:CBRC7TM_97.
ACCESSION	AB065534
VERSION	AB065534.1
KEYWORDS	GI:21928364
SOURCE	Homo sapiens (human)
ORGANISM	Homo sapiens
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
AUTHORS	1 Suwa,M., Sato,T., Okouchi,I., Arita,M., Futami,K., Matsumoto,S., Tsutsumi,S., Aburatani,H., Asai,K. and Akiyama,Y.
TITLE	Genome-wide discovery and analysis of human seven transmembrane helix receptor genes
JOURNAL	Unpublished
REFERENCE	2 (bases 1 to 1390)
AUTHORS	suwa,M.
TITLE	Direct Submission
JOURNAL	Submitted (11-JUN-2001) Makiko Suwa Computational Biology Department

COMMENT

Tel: 81-3-3599-8080, Fax: 81-3-3599-8081

This sequence is a seven transmembrane helix receptor candidate predicted from the whole human genome sequences using our automated system that contains programs of Gene finding(GeneDecoder), sequence search, motif-domain assignment and transmembrane helix prediction.

And the sequence is submitted by the collaborative project between [Computational Biology Research Center (CBRC), National Institute of Advanced Industrial Science and Technology (AIST)] and [Genome Science Division, Research Center for Advanced Science and Technology (RCAST), University of Tokyo].

FEATURES
source

```

Location/Qualifiers
1. .1190
  /organism="Homo sapiens"
  /mol_type="genomic DNA"
  /isolate="CBK71M_97"
  /db_xref="taxon:9606"
  /chromosome="11"
  201. .1190
  /codon_start=1
  /evidence=not experimental
  /product="seven transmembrane helix receptor"
  /protein_id="BAC05780.1"

```

/db_xref="GI:21928365"
/translation="MSSTLGHNMESPNHVDVPSVFFLLGIPGLEQPHLMLSLPVCGL
GTATVGNITLVVATEPVLHKEVYFLCLMLSTVDLAASVSTVPKLLAIHFWCAGHI
SASACLAQMPFIHAFNMESVLLAMAFDRVAICHLRYATILITDIIAHIGVAAY
RGSLLMLPCPLIGLNFQSHVILHTYCEHMAVVKLACGDRNRYVGLTAALLVIG
VDLFCIGLSYALSQAQVLRLLSSHEARSKALGTGSHVVCVILISYTPALFSPFTRFHG
HVPVHIHILLANVILLPPALNPVYGVKTKQIRKRVVRVFSQSGMGIKASE"

ORIGIN

Query Match	99.4%;	Score	980.6;	DB	9;	Length	1390;
Best Local Similarity	99.6%;	Pred. No.	1.4e-286;				
Matches	983;	Conservative	0;	Mismatches	4;	Indels	0;
Gaps	0;						
QY	1	ATGTCACAGCTCTTGGCCACAAATGGAATCTCTCATCACTGATGTTGACCTTCT	60				
DB	201	ATGTCACAGCTCTTGGCCACAAATGGAATCTCTTAATCACTGATGTTGACCTTCT	260				
QY	61	GTCTTCTTCTCTGGGCAATCCAGTCTGGAACAATTTCAATTTGCGCTCTCACTCCCT	120				
DB	261	GTCTTCTTCTCTGGGCAATCCAGGCTCTGGAACAATTTCAATTTGTTGGCTCTCACTCCCT	320				
QY	121	GTGTGTGGCTTAGGCACAGCCACAAATGTGGGCAATATAACTATTCTGTGTGTTGTGCG	180				
DB	321	GTGTGTGGCTTAGGCACAGCCACAAATGTGGGCAATATAACTATTCTGTGTGTTGTGCG	380				
QY	181	ACTGAACAGCTTTGCAACAGCCTGTGTACCTTTTCTGTGATGCTCTCAACATCGAC	240				
DB	381	ACTGAACAGCTTTGCAACAGCCTGTGTACCTTTTCTGTGATGCTCTCAACATCGAC	440				
QY	241	TTGGCTGCTCTGTCTCCACAGTCTCCCAAGCTACTGGCTATCTTCTGGTGGAGCCGA	300				
DB	441	TTGGCTGCTCTGTCTCCACAGTCTCCCAAGCTACTGGCTATCTTCTGGTGGAGCCGA	500				
QY	301	CATATATCTGCTCTGCTGCTGCGACATATGTTCTTCATTCATGCTCTTGCATGATG	360				
DB	501	CATATATCTGCTCTGCTGCTGCGACAGATGTTCTTCATTCATGCTCTTGCATGATG	560				
QY	361	GAGTCAGCTGTGCTACTGGCCATGCGCTTTGATCGTAGTGGCCATCTGCCACCCATC	420				
DB	561	GAGTCAGCTGTGCTACTGGCCATGCGCTTTGATCGTAGTGGCCATCTGCCACCCATC	620				
QY	421	CGCTATGCCAATCTCTCACTGACACCATCATTTGCCACATAGGGGTGGCAGCTGATG	480				
DB	621	CGCTATGCCAATCTCTCACTGACACCATCATTTGCCACATAGGGGTGGCAGCTGATG	680				
QY	481	CGAGGCTCCCTGCTCATGCTCCCATGTCCTTCTTTATTTGGGCGTTTGAATCTTCCAA	540				
DB	681	CGAGGCTCCCTGCTCATGCTCCCATGTCCTTCTTTATTTGGGCGTTTGAATCTTCCAA	740				
QY	541	AGCCATGTGATCTTACACAGCTACTGTGAGCACATGCTGTGGTGAAGCTGGCTGTGGA	600				
DB	741	AGCCATGTGATCTTACACAGCTACTGTGAGCACATGCTGTGGTGAAGCTGGCTGTGGA	800				
QY	601	GACACAGGCTTAACCGTGTGTATGGGCTGACAGCTGCACTGTTGGTGTGATTTGGGTTGAC	660				
DB	801	GACACAGGCTTAACCGTGTGTATGGGCTGACAGCTGCACTGTTGGTGTGATTTGGGTTGAC	860				
QY	661	TTGTTTGTGATTTGGTCTCTCTATGCGCTTAATTTGCACAAGCTGTGCTTGCCTCTCATCC	720				
DB	861	TTGTTTGTGATTTGGTCTCTCTATGCGCTTAAGTGCACAAGCTGTGCTTGCCTCTCATCC	920				
QY	721	CATGAAGCTGGTCCAAAGGCTTAGGACCTGTGGTTTCCCATGCTGTGTCACTCTCATC	780				
DB	921	CATGAAGCTGGTCCAAAGGCTTAGGACCTGTGGTTTCCCATGCTGTGTCACTCTCATC	980				
QY	781	TCTTATACACAGGCTCTTCTCTCTTTTATACACACCGCTTTGGCCATCAGTTCAGTC	840				
DB	981	TCTTATACACAGGCTCTTCTCTCTTTTATACACACCGCTTTGGCCATCAGTTCAGTC	1040				
QY	841	CATATTACATCTTTTGGCCAAATGTTATCTGCTTTTGGCCACTGCTCTTAATCTGTG	900				
DB	1041	CATATTACATCTTTTGGCCAAATGTTATCTGCTTTTGGCCACTGCTCTTAATCTGTG	1100				

QY	901	GTATATGAGTTAAGACCAACACATCGTAAAGAGTTGTCAAGGTGTTTCAAGTGGG	960
DB	1101	GTATATGAGTTAAGACCAACACATCGTAAAGAGTTGTCAAGGTGTTTCAAGTGGG	1160
QY	961	CAGGAATGGGCATCAAGGCATCTGAG	987
DB	1161	CAGGAATGGGCATCAAGGCATCTGAG	1187

RESULT 8

AC026090/c
LOCUS
DEFINITION
AC026090
Homo sapiens chromosome 11 clone RP11-658K18, WORKING DRAFT
SEQUENCE, 19 unordered pieces.
AC026090
AC026090.5 GI:9958310
HTG; HTGS PHASE1; HTGS_DRAFT.
KEYWORDS
SOURCE
Homo sapiens
Homo sapiens
ORGANISM
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
1 (bases 1 to 172027)
Waterston,R.H.
The sequence of Homo sapiens clone
Unpublished
2 (bases 1 to 172027)
Waterston,R.H.
Direct Submission
Submitted (19-MAR-2000) Genome Sequencing Center, Washington
University School of Medicine, 4444 Forest Park Parkway, St. Louis,
MO 63108, USA
COMMENT
On Sep 1, 2000 this sequence version replaced gi:9838305.

----- Genome Center -----
Center: Washington University Genome Sequencing Center
Center code: WUGSC
Web site: http://genome.wustl.edu/gsc/index.shtml
----- Project Information -----
Center project name: H NH0658K18
----- Summary Statistics -----
Sequencing vector: M13; 100%
Sequencing method: plasmid; 0%
Chemistry: Dye-terminator Big Dye; 0% of reads
Assembly program: Phrap; version 0.990319
Consensus quality: 160996 bases at least Q40
Consensus quality: 163968 bases at least Q30
Consensus quality: 165626 bases at least Q20
Insert size: 180000; agarose-fp
Insert size: 171387; sum-of-contigs
Quality coverage: 4.33 in Q20 bases; agarose-fp
Quality coverage: 4.82 in Q20 bases; sum-of-contigs

----- NOTE: This is a 'working draft' sequence. It currently
* consists of 19 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.

1	1526:	contig of 1526 bp in length
*	1527	1626: gap of unknown length
*	1627	2975: contig of 1349 bp in length
*	2376	3075: gap of unknown length
*	3076	5073: contig of 1998 bp in length
*	5074	5173: gap of unknown length
*	5174	9317: contig of 4144 bp in length
*	9318	9417: gap of unknown length
*	9418	13195: contig of 3778 bp in length
*	13196	13295: gap of unknown length
*	13296	16482: contig of 3187 bp in length
*	16483	16582: gap of unknown length

16583 21965: contig of 5383 bp in length
* 21966 gap of unknown length
* 22065: contig of 11102 bp in length
* 33167: contig of unknown length
* 33267: gap of 9895 bp in length
* 43162: contig of 9895 bp in length
* 43262: gap of unknown length
* 53157: contig of 9895 bp in length
* 53257: gap of unknown length
* 61965: contig of 8708 bp in length
* 62065: gap of unknown length
* 70847: contig of 8782 bp in length
* 70947: gap of unknown length
* 84297: contig of 13350 bp in length
* 84397: gap of unknown length
* 84398 gap of unknown length
* 100019: contig of 15622 bp in length
* 100119: gap of unknown length
* 114785: contig of 14666 bp in length
* 114885: gap of unknown length
* 131872: contig of 16987 bp in length
* 131973: gap of unknown length
* 131973: contig of 15646 bp in length
* 147118: gap of unknown length
* 171498: contig of 23780 bp in length
* 171499 gap of unknown length
* 171598: contig of 429 bp in length.
* 171599

FEATURES

source

Location/Qualifiers

1. 172027
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/chromosome="11"
/clone="RP11-658K18"
1. 1526
/note="assembly_name:Contig14"
misc_feature 1627..2975
/note="assembly_name:Contig17"
misc_feature 3076..5073
/note="assembly_name:Contig18"
misc_feature 5174..9317
/note="assembly_name:Contig19"
misc_feature 9418..13195
/note="assembly_name:Contig20"
misc_feature 13296..16482
/note="assembly_name:Contig21"
misc_feature 16583..21965
/note="assembly_name:Contig22"
misc_feature 22066..33167
/note="assembly_name:Contig23"
misc_feature 33268..43162
/note="assembly_name:Contig24"
misc_feature 43263..53157
/note="assembly_name:Contig25"
misc_feature 53258..61965
/note="assembly_name:Contig26"
misc_feature 62066..70847
/note="assembly_name:Contig27
clone_end:T7
vector_side:right"
misc_feature 70948..84297
/note="assembly_name:Contig28"
misc_feature 84398..100019
/note="assembly_name:Contig29"
misc_feature 100120..114785
/note="assembly_name:Contig30"
misc_feature 114886..131872
/note="assembly_name:Contig31"
misc_feature 131973..147618
/note="assembly_name:Contig32"
misc_feature 147719..171498
/note="assembly_name:Contig33"
misc_feature 171599..172027
/note="assembly_name:Contig16"

ORIGIN

Query Match

Best Local Similarity 99.4%; Score 980.6; DB 2; Length 172027;

Matches 983; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 ATGTCAGACACTCTTGGCCACACATGGAATCTCTCATCAGCTGATGTGACCCCTTCT 60
Db 89467 ATGTCAGACACTCTTGGCCACACATGGAATCTCTCATCAGCTGATGTGACCCCTTCT 89408
QY 61 GTCTTCTTCTCTCTGGGCATCCAGGCTGGAACATTTTCATTTGGCTCTCAGTCCCT 120
Db 89407 GTCTTCTTCTCTGGGCATCCAGGCTGGAACATTTTCATTTGGCTCTCAGTCCCT 89348
QY 121 GTGTGGCTTAGGCACAGCCAAATTTGGGCAATATAACTATTCTGGTGTGTGTGCC 180
Db 89347 GTGTGGCTTAGGCACAGCCAAATTTGGGCAATATAACTATTCTGGTGTGTGTGCC 89288
QY 181 ACTGAACCAAGTCTTGACAAAGCTGTGTACCTTTTCTGTGATGCTCTCAACCATCGAC 240
Db 89287 ACTGAACCAAGTCTTGACAAAGCTGTGTACCTTTTCTGTGATGCTCTCAACCATCGAC 89228
QY 241 TTGGCTGCTCTGTCTCCACAGTTCCTCAAGCTACTGGCTATCTCTGTGTGGAGCCGGA 300
Db 89227 TTGGCTGCTCTGTCTCCACAGTTCCTCAAGCTACTGGCTATCTCTGTGTGGAGCCGGA 89168
QY 301 CATATATCTGCTCTGCTGCTGCTGGCACATATATGTTCTTCAATCATGCCCTTCTGCATGATG 360
Db 89167 CATATATCTGCTCTGCTGCTGCTGGCACATATATGTTCTTCAATCATGCCCTTCTGCATGATG 89108
QY 361 GAGTCCACTGTGCTACTGGCCATGSCCTTTGATCGCTAGTGGCCATCTGCCACCATCTC 420
Db 89107 GAGTCCACTGTGCTACTGGCCATGSCCTTTGATCGCTAGTGGCCATCTGCCACCATCTC 89048
QY 421 CGCTATGCCACAATCCTCCTGACCATCATTTGCCACATAGGCGTGGAGCTGTAGTG 480
Db 89047 CGCTATGCCACAATCCTCCTGACCATCATTTGCCACATAGGCGTGGAGCTGTAGTG 88988
QY 481 CGAGGCTCCCTGCTCAAGCTCCCATGCTCCCTTTTATTGGGCGTTTGAACCTTCTGCCAA 540
Db 88987 CGAGGCTCCCTGCTCAAGCTCCCATGCTCCCTTTTATTGGGCGTTTGAACCTTCTGCCAA 88928
QY 541 AGCCATGTGATCCTACACACGCTACTGTGAGCACATGCTGTGGTGAAGCTGCGCTTGA 600
Db 88927 AGCCATGTGATCCTACACACGCTACTGTGAGCACATGCTGTGGTGAAGCTGCGCTTGA 88868
QY 601 GACACACAGGCTTAACCGGTGTATGGGCTGACAGCTGACCTGTTGGTCAATTTGGGTTGAC 660
Db 88867 GACACACAGGCTTAACCGGTGTATGGGCTGACAGCTGACCTGTTGGTCAATTTGGGTTGAC 88808
QY 661 TTGTTTGGCATTTGCTCTCTCTATGCCCTAAATTCACAAAGCTGCTTCGCTCTCTCATCC 720
Db 88807 TTGTTTGGCATTTGCTCTCTCTATGCCCTAAATTCACAAAGCTGCTTCGCTCTCTCATCC 88748
QY 721 CATGAAGCTGGTCCAAAGGCTTAGGACCTGTGGTCCCATGCTGTGTGATCTCATCTCATC 780
Db 88747 CATGAAGCTGGTCCAAAGGCTTAGGACCTGTGGTCCCATGCTGTGTGATCTCATCTCATC 88688
QY 781 TCTTATACACAGGCTCTCTCTCTCTTTTACACACGCTTTGGCCATCAGTTCCAGTC 840
Db 88687 TCTTATACACAGGCTCTCTCTCTTTTACACACGCTTTGGCCATCAGTTCCAGTC 88628
QY 841 CATATTCAATCTTTTGGCCAAATGTTTATCTGCTTTTGGCACCTGCTCTTAATCTGTG 900
Db 88627 CATATTCAATCTTTTGGCCAAATGTTTATCTGCTTTTGGCACCTGCTCTTAATCTGTG 88568
QY 901 GTATATCGAGTTAAGACCAACAGATCCGTAAGAAGTTGTCAAGGTGTTTCAAAGTGGG 960
Db 88567 GTATATCGAGTTAAGACCAACAGATCCGTAAGAAGTTGTCAAGGTGTTTCAAAGTGGG 88508
QY 961 CAGGGAATGGGCATCAAGGCATCTGAG 987
Db 88507 CAGGGAATGGGCATCAAGGCATCTGAG 88481

RESULT 9
AC116156/c
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS
TITLE
JOURNAL
REFERENCE
AUTHORS

AC116156 176597 bp DNA linear PRI 13-MAY-2002
 Homo sapiens chromosome 11, clone RP11-44D14, complete sequence.
 AC116156
 AC116156.3 GI:20522208
 HTG.
 Homo sapiens (human)
 Homo sapiens
 Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 1 (bases 1 to 176597)
 Birren,B., Linton,L., Nusbaum,C. and Lander,E.
 Homo sapiens chromosome 11, clone RP11-44D14
 Unpublished
 2 (bases 1 to 176597)
 Birren,B., Linton,L., Nusbaum,C., Lander,E., Ali,A., Allen,N.,
 Anderson,S., Barna,N., Bastien,V., Bloom,T., Boguslavskiy,L.,
 Boukhgalter,B., Brown,A., Camarata,J., Campopiano,A., Chang,J.,
 Chazaro,B., Choepel,Y., Colangelo,M., Collins,S., Collymore,A.,
 Cook,A., Cooke,P., DeArelano,K., Dewar,K., Diaz,J.S., Dodge,S.,
 Ginde,S., Gord,S., Goyette,M., Graham,L., Grand-Pierre,N.,
 Hagoes,B., Horton,L., Hulme,W., Iliev,I., Johnson,R., Jones,C.,
 Kamat,A., Karatas,A., Kells,C., LaRocque,K., Lamazares,R.,
 Landers,T., Lehoczyk,J., Levine,R., Lindblad-Toh,K., Liu,G.,
 MacLean,C., Macdonald,P., Major,J., Marquis,N., Matthews,C.,
 McCarthy,M., McEwan,P., McKernan,K., Meldrum,J., Meneus,L.,
 Mihova,T., Mlenga,V., Murphy,T., Naylor,J., Nguyen,C., Nicol,R.,
 Norbu,C., Norman,C.H., O'Connor,T., O'Donnell,P., O'Neil,D.,
 Oliver,J., Peterson,K., Phunkhang,P., Pierre,N., Pollara,V.,
 Raymond,C., Retta,R., Rieback,M., Riley,R., Rise,C., Rogov,P.,
 Roman,J., Rosetti,M., Roy,A., Santos,R., Schauer,S., Schupback,R.,
 Seaman,S., Severy,P., Spencer,B., Stange-Thomann,N., Stojanovic,N.,
 Strauss,N., Subramanian,A., Talamas,J., Tesfaye,S., Theodore,J.,
 Topham,K., Travers,M., Travis,N., Trigilio,J., Vassiliev,H.,
 Viel,R., Vo,A., Wilson,B., Wu,X., Wyman,D., Ye,W.J., Young,G.,
 Zainoun,J., Zembek,L., Zimmer,A. and Zody,M.
 Direct Submission
 Submitted (25-MAR-2002) Whitehead Institute/MIT Center for Genome
 Research, 320 Charles Street, Cambridge, MA 02141, USA
 3 (bases 1 to 176597)

TITLE
JOURNAL
REFERENCE
AUTHORS

Birren,B., Linton,L., Nusbaum,C., Lander,E., Ali,A., Allen,N.,
 Anderson,S., Barna,N., Bastien,V., Bloom,T., Boguslavskiy,L.,
 Boukhgalter,B., Brown,A., Camarata,J., Campopiano,A., Chang,J.,
 Chazaro,B., Choepel,Y., Colangelo,M., Collins,S., Collymore,A.,
 Cook,A., Cooke,P., DeArelano,K., Dewar,K., Diaz,J.S., Dodge,S.,
 Ginde,S., Gord,S., Goyette,M., Graham,L., Grand-Pierre,N.,
 Hagoes,B., Horton,L., Hulme,W., Iliev,I., Johnson,R., Jones,C.,
 Kamat,A., Karatas,A., Kells,C., LaRocque,K., Lamazares,R.,
 Landers,T., Lehoczyk,J., Levine,R., Lindblad-Toh,K., Liu,G.,
 MacLean,C., Macdonald,P., Major,J., Marquis,N., Matthews,C.,
 McCarthy,M., McEwan,P., McKernan,K., Meldrum,J., Meneus,L.,
 Mihova,T., Mlenga,V., Murphy,T., Naylor,J., Nguyen,C., Nicol,R.,
 Norbu,C., Norman,C.H., O'Connor,T., O'Donnell,P., O'Neil,D.,
 Oliver,J., Peterson,K., Phunkhang,P., Pierre,N., Pollara,V.,
 Raymond,C., Retta,R., Rieback,M., Riley,R., Rise,C., Rogov,P.,
 Roman,J., Rosetti,M., Roy,A., Santos,R., Schauer,S., Schupback,R.,
 Seaman,S., Severy,P., Spencer,B., Stange-Thomann,N., Stojanovic,N.,
 Strauss,N., Subramanian,A., Talamas,J., Tesfaye,S., Theodore,J.,
 Topham,K., Travers,M., Travis,N., Trigilio,J., Vassiliev,H.,
 Viel,R., Vo,A., Wilson,B., Wu,X., Wyman,D., Ye,W.J., Young,G.,
 Zainoun,J., Zembek,L., Zimmer,A. and Zody,M.
 Direct Submission
 Submitted (25-MAR-2002) Whitehead Institute/MIT Center for Genome
 Research, 320 Charles Street, Cambridge, MA 02141, USA
 3 (bases 1 to 176597)

TITLE
JOURNAL
REFERENCE
AUTHORS

Birren,B., Linton,L., Nusbaum,C., Lander,E., Ali,A., Allen,N.,
 Anderson,S., Barna,N., Bastien,V., Bloom,T., Boguslavskiy,L.,
 Boukhgalter,B., Brown,A., Camarata,J., Campopiano,A., Chang,J.,
 Chazaro,B., Choepel,Y., Colangelo,M., Collins,S., Collymore,A.,
 Cook,A., Cooke,P., DeArelano,K., Dewar,K., Diaz,J.S., Dodge,S.,
 Ginde,S., Gord,S., Goyette,M., Graham,L., Grand-Pierre,N.,
 Hagoes,B., Horton,L., Hulme,W., Iliev,I., Johnson,R., Jones,C.,
 Kamat,A., Karatas,A., Kells,C., LaRocque,K., Lamazares,R.,
 Landers,T., Lehoczyk,J., Levine,R., Lindblad-Toh,K., Liu,G.,
 MacLean,C., Macdonald,P., Major,J., Marquis,N., Matthews,C.,
 McCarthy,M., McEwan,P., McKernan,K., Meldrum,J., Meneus,L.,
 Mihova,T., Mlenga,V., Murphy,T., Naylor,J., Nguyen,C., Nicol,R.,
 Norbu,C., Norman,C.H., O'Connor,T., O'Donnell,P., O'Neil,D.,
 Oliver,J., Peterson,K., Phunkhang,P., Pierre,N., Pollara,V.,
 Raymond,C., Retta,R., Rieback,M., Riley,R., Rise,C., Rogov,P.,
 Roman,J., Rosetti,M., Roy,A., Santos,R., Schauer,S., Schupback,R.,
 Seaman,S., Severy,P., Spencer,B., Stange-Thomann,N., Stojanovic,N.,
 Strauss,N., Subramanian,A., Talamas,J., Tesfaye,S., Theodore,J.,
 Topham,K., Travers,M., Travis,N., Trigilio,J., Vassiliev,H.,
 Viel,R., Vo,A., Wilson,B., Wu,X., Wyman,D., Ye,W.J., Young,G.,
 Zainoun,J., Zembek,L., Zimmer,A. and Zody,M.
 Direct Submission
 Submitted (13-MAY-2002) Whitehead Institute/MIT Center for Genome
 Research, 320 Charles Street, Cambridge, MA 02141, USA
 On May 10, 2002 this sequence version replaced gi:2017753.
 All repeats were identified using RepeatMasker:
 Smit, A.F.A. & Green, P. (1996-1997)
 http://ftp.genome.washington.edu/RM/RepeatMasker.html

TITLE
JOURNAL
REFERENCE
AUTHORS

Center: Whitehead Institute/ MIT Center for Genome Research
 Center code: WIBR
 Web site: http://www-seq.wi.mit.edu
 Contact: sequence_submissions@genome.wi.mit.edu
 ----- Project Information
 Center project name: L26176
 Center clone name: 44_D_14
 ----- Location/Qualifiers
 1. .176597
 /organism="Homo sapiens"
 /mol_type="genomic DNA"
 /db_xref="taxon:9606"
 /chromosome="11"
 /map="11"
 /clone="RP11-44D14"
 /clone_lib="RP11-11 Human Male BAC"
 775..802
 repeat_region
 /rpt_family="(TTCC)n"
 1179..1224
 repeat_region
 /rpt_family="(CA)n"

QY 565 TGTGAGCATGGCTGTGGTGAAGCTGGCTGTGGAGACACAGCCCTAAACCGTGTGTAT 624
 Db |||||
 QY 541 TGTGAGCATGGCTGTGGTGAAGCTGGCTGTGGAGACACAGCCCTAAACCGTGTGTAT 600
 Db |||||
 QY 625 GGGCTGACAGCTGCACTGTGGTCAATGGGGTTGACTTGTGTTGCAATGGTCTCTCCAT 684
 Db |||||
 QY 601 GGGCTGACAGCTGCACTGTGGTCAATGGGGTTGACTTGTGTTGCAATGGTCTCTCCAT 660
 Db |||||
 QY 685 GCCCTAATGACAGCTGCTTGGCTCTCAATCCCATGAAGCTCGGTCCAAAGCCCTTA 744
 Db |||||
 QY 661 GCCCTAAGTGACAGCTGCTTGGCTCTCAATCCCATGAAGCTCGGTCCAAAGCCCTTA 720
 Db |||||
 QY 745 GGGACCTGTGGTCCCATGTCTGGTCAATGGGGTTGACTTGTGTTGCAATGGTCTCTCC 804
 Db |||||
 QY 721 GGGACCTGTGGTCCCATGTCTGGTCAATGGGGTTGACTTGTGTTGCAATGGTCTCTCC 780
 Db |||||
 QY 805 TTTTITACACACCGCTTTGGCCATCACGTTCCAGTCCCATATTCACATTTTGGCCAAAT 864
 Db |||||
 QY 781 TTTTITACACACCGCTTTGGCCATCACGTTCCAGTCCCATATTCACATTTTGGCCAAAT 840
 Db |||||
 QY 865 GTTATCTGCTTTGGCACTGCTCTTAATCTGTGGTATATGAGTTAAGACCAACAG 924
 Db |||||
 QY 841 GTTATCTGCTTTGGCACTGCTCTTAATCTGTGGTATATGAGTTAAGACCAACAG 900
 Db |||||
 QY 925 ATCCGTAAGAGTTGTCAGGCTGTTTCAAGTGGCAGGGAATGGGCATCAAGGCATCT 984
 Db |||||
 QY 901 ATCCGTAAGAGTTGTCAGGCTGTTTCAAGTGGCAGGGAATGGGCATCAAGGCATCT 960
 Db |||||
 QY 985 GAG 987
 Db |||||
 QY 961 GAG 963

RESULT 12
 AX448757
 LOCUS
 DEFINITION Sequence 423 from Patent WO0224726.
 ACCESSION AX448757
 VERSION AX448757.1 GI:21697655
 KEYWORDS
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens

REFERENCE
 AUTHORS Veithen, A.
 TITLE Olfactory and pheromones g-protein coupled receptors
 JOURNAL Patent: WO 0224726-A 423 28-MAR-2002;
 ChemCom S.A. (BE)

FEATURES
 source Location/Qualifiers
 1..966
 /organism="Homo sapiens"
 /mol_type="unassigned DNA"
 /db_xref="taxon:9606"
 1..966
 /note="unnamed protein product"
 /codon_start=1
 /protein_id="CAD37694.1"
 /db_xref="GI:21697655"
 /db_xref="REMTREMBL:CAD37694"
 /translation="MESPNHTDVPSPVFFLLGIPGLPQLWLSLPVCGLTATVGN
 ITILVVATPVLHPVLPFLQMLSTDLAASVTVFKLLAIWFGAGHISASACLAQ
 MFFIHCMMESTVLLAMADRYVAICPLRYATILITDIIAHIGVAAVVRGSLMLP
 CPELIGRLNFCQSHVLIHTCEHMAVVKACGDPTRNVYGLTAALIVIGVDFICGL
 SYALSAQVLRSLSHSARSKALGTGSHVCVILISYTPALFSFTFHRFGHVPVHIH
 LLANVYLLPALNPVYVGVKTKQIRKRVVRVFSQGGMGIKASE"

CDS

Query Match 96.9%; Score 956.6; DB 6; Length 966;
 Best Local Similarity 99.6%; Pred. No. 2.7e-279;
 Matches 959; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

ORIGIN

RESULT 13
 AC109341/c
 LOCUS
 DEFINITION Homo sapiens chromosome 11, clone CTD-2504M7, complete sequence.

QY 25 ATGGAATCTCCTCATCACATGATGTTGAACCTTCTGTCTTCTCCTCTGGGCAATCCA 84
 Db |||||
 QY 1 ATGGAATCTCCTTAATCACATGATGTTGAACCTTCTGTCTTCTCCTCTGGGCAATCCA 60
 Db |||||
 QY 85 GGTCTGGAAACAATTTCAATTTGGCTCTCACTCCCTGTGTGGCTTAGGCACACGCCACA 144
 Db |||||
 QY 61 GGTCTGGAAACAATTTCAATTTGGCTCTCACTCCCTGTGTGGCTTAGGCACACGCCACA 120
 Db |||||
 QY 145 ATTGTGGGCAATAAATAATTTCTGTTGGCACTGAACCAAGCTTTGTCACAAGCT 204
 Db |||||
 QY 121 ATTGTGGGCAATAAATAATTTCTGTTGGTCTGTCACCTGAACCAAGCTTTGTCACAAGCT 180
 Db |||||
 QY 205 GTGTACCTTTTCTGTGCAATGCTCTCAACCAATGACCTTGGCTGTGTCTGTCTCCACAGTT 264
 Db |||||
 QY 181 GTGTACCTTTTCTGTGCAATGCTCTCAACCAATGACCTTGGCTGTGTCTGTCTCCACAGTT 240
 Db |||||
 QY 265 CCCAAGCTTACTGGCTATCTTCTGGTGTGGAGCCGACATATATCTGCTCTGCTGCTGCTG 324
 Db |||||
 QY 241 CCCAAGCTTACTGGCTATCTTCTGGTGTGGAGCCGACATATATCTGCTCTGCTGCTGCTG 300
 Db |||||
 QY 325 GCACATATGTTTCTTCAATTCATGCTTCTGCAATGATGAGTCCACTGTGCTACTGTGCCATG 384
 Db |||||
 QY 301 GCACATATGTTTCTTCAATTCATGCTTCTGCAATGATGAGTCCACTGTGCTACTGTGCCATG 360
 Db |||||
 QY 385 GCCTTTGATCGCTACGTGGCCATCTGCCACCCTCGCTATGCGCAATCTCTCACTGAC 444
 Db |||||
 QY 361 GCCTTTGATCGCTACGTGGCCATCTGCCACCCTCGCTATGCGCAATCTCTCACTGAC 420
 Db |||||
 QY 445 ACCATCATTTGCCACATAGGGGTGGCAGCTGTAGTGGAGGCTCCCTGCTCATGTGCCA 504
 Db |||||
 QY 421 ACCATCATTTGCCACATAGGGGTGGCAGCTGTAGTGGAGGCTCCCTGCTCATGTGCCA 480
 Db |||||
 QY 505 TGTCCCTTCTTTAATTTGGGCGTTTGAACCTTCTGCCAAAGCCAATGTGATCTTACACACGTAC 564
 Db |||||
 QY 481 TGTCCCTTCTTTAATTTGGGCGTTTGAACCTTCTGCCAAAGCCAATGTGATCTTACACACGTAC 540
 Db |||||
 QY 565 TGTGAGCAGATGGCTGTGGTGAAGCTGGCTGTGGAGACACAGGCCCTAAACCGTGTGTAT 624
 Db |||||
 QY 541 TGTGAGCAGATGGCTGTGGTGAAGCTGGCTGTGGAGACACAGGCCCTAAACCGTGTGTAT 600
 Db |||||
 QY 625 GGGCTGACAGCTGCATGTTGGTCAATTTGGGTTGACTTGTGTTTGCATTTGGTCTCTCCTAT 684
 Db |||||
 QY 601 GGGCTGACAGCTGCATGTTGGTCAATTTGGGTTGACTTGTGTTTGCATTTGGTCTCTCCTAT 660
 Db |||||
 QY 685 GCCCTAATGACAAAGCTGCTTTCGCTCTCATFCCCATGAAGCTCGGTCCAAAGCCCTTA 744
 Db |||||
 QY 661 GCCCTAATGACAAAGCTGCTTTCGCTCTCATFCCCATGAAGCTCGGTCCAAAGCCCTTA 720
 Db |||||
 QY 745 GGGACCTGTGGTCCCATGTCTGTGTCATCTCTTATACACAGCCCTCTTCTCC 804
 Db |||||
 QY 721 GGGACCTGTGGTCCCATGTCTGTGTCATCTCTTATACACAGCCCTCTTCTCC 780
 Db |||||
 QY 805 TTTTITACACACCGCTTTGGCCATCACGTTCCAGTCCCATATTCACATTTTGGCCAAAT 864
 Db |||||
 QY 781 TTTTITACACACCGCTTTGGCCATCACGTTCCAGTCCCATATTCACATTTTGGCCAAAT 840
 Db |||||
 QY 865 GTTATCTGCTTTTGGCACTGCTCTTAATCTGTGGTATATGAGTTAAGACCAACAG 924
 Db |||||
 QY 841 GTTATCTGCTTTTGGCACTGCTCTTAATCTGTGGTATATGAGTTAAGACCAACAG 900
 Db |||||
 QY 925 ATCCGTAAGAGTTGTCAGGCTGTTTCAAGTGGCAGGGAATGGGCATCAAGGCATCT 984
 Db |||||
 QY 901 ATCCGTAAGAGTTGTCAGGCTGTTTCAAGTGGCAGGGAATGGGCATCAAGGCATCT 960
 Db |||||
 QY 985 GAG 987
 Db |||||
 QY 961 GAG 963

RESULT 13
 AC109341/c
 LOCUS
 DEFINITION Homo sapiens chromosome 11, clone CTD-2504M7, complete sequence.


```
/rpt family="HAL1"
complement(13044..13443)
/rpt family="LiMA9"
complement(13639..13939)
/rpt family="AluSx"
16194..16222
/rpt family="AT_rich"
complement(17502..17712)
/rpt family="LiMA9"
complement(17688..18018)
/rpt family="LiMA9"
complement(18247..18311)
/rpt family="L2"
complement(19166..19280)
/rpt family="MIR"
19450..19637
/rpt family="MIR"
21072..21116
/rpt family="LiMB1"
complement(21146..21534)
/rpt family="LiMA4"
21536..21774
/rpt family="LiMA4"
22101..22162
/rpt family="(TTCA)n"
23080..23388
/rpt family="AluSq"
23817..23955
/rpt family="Charliel"
complement(24081..24150)
/rpt family="LiMB3"
complement(24151..24251)
/rpt family="LiMA1"
24254..24829
/rpt family="LiMA1"
complement(24841..25271)
/rpt family="LiMB6"
25424..25503
/rpt family="(TA)n"
25712..26331
/rpt family="AT_rich"
26719..26754
/rpt family="AT_rich"
27022..27136
/rpt family="(TTATA)n"
28928..28957
/rpt family="AT_rich"
29743..30367
/rpt family="LiPB3"
complement(30373..30691)
/rpt family="LiMB6"
complement(30926..30987)
```

Query Match 96.6%; Score 953.2; DB 9; Length 202761;

Best Local Similarity 99.2%; Pred. No. 5.4e-278;

Matches 980; Conservative 0; Mismatches 3; Indels 5; Gaps 2;

QY 1 ATGTCCAGCACTCTGGCCACCAACATGGAATCTCTCATCACACTGATGTTGACCCCTTC- 59

Db 69631 ATGTCCAGCACTCTGGCCACCAACATGGAATCTCTCATCACACTGATGTTGACCCCTTC 69572

QY 60 TGTCTTCTCTCTCTGGGATCCAGCTCGGACAAATTTTCAATTTGGGCTCTCACTCCC 119

Db 69571 TGTCTTCTCTCTCTGGGATCCAGCTCGGACAAATTTTCAATTTGGGCTCTCACTCCC 69512

QY 120 TGTGTGGCTTAGGCAGCAGCAAAATTTGGGCAATATACTATTCTGGTGTGTTGTC 179

Db 69511 TGTGTGGCTTAGGCAGCAGCAAAATTTGGGCAATATACTATTCTGGTGTGTTGTC 69452

QY 180 CACTGAACCACTCTTGACAAAGCTGTGTACCTTTTCTGTGCAATGCTTCAACCATCGA 239

Db 69451 CACTGAACCACTCTTGACAAAGCTGTGTACCTTTTCTGTGCAATGCTTCAACCATCGA 69392

```
QY 240 CTTGGTGCCTCTGTCTCCACAGTTCCCAAGCTACTGGCTATCTTCTGGTGTGAGCCGG 299
Db 69391 CTTGGTGCCTCTGTCTCCACAGTTCCCAAGCTACTGGCTATCTTCTGGTGTGAGCCGG 69332
QY 300 ACATATATCTGCCTCTGCCTGGGACATATGTTCTTCAATTCATGCTTCTGTCATGAT 359
Db 69331 ACATATATCTGCCTCTGCCTGGGACATATGTTCTTCAATTCATGCTTCTGTCATGAT 69272
QY 360 GGAGTCCACTGTGTACTGGCCATGGCCTTTTATCGCTACGTGGCCATCTGCGACCCACT 419
Db 69271 GGAGTCCACTGTGTACTGGCCATGGCCTTTTATCGCTACGTGGCCATCTGCGACCCACT 69212
QY 420 CCGCTATGCCACAATCTCACTGACACCATATTTGCCACATAGGGGTGGAGCTGTAGT 479
Db 69211 CCGCTATGCCACAATCTCACTGACACCATATTTGCCACATAGGGGTGGAGCTGTAGT 69152
QY 480 CGGAGGCTCCCTGTCTATGCTCCCATGTCCTTCTTATTTGGGGTGTGAACTTCTGCCA 539
Db 69151 CGGAGGCTCCCTGTCTATGCTCCCATGTCCTTCTTATTTGGGGTGTGAACTTCTGCCA 69092
QY 540 AAGCCATGTGATCTCTACACAGTACTGTGAGCACATGCTGTGGTGAAGCTGGCTGTGG 599
Db 69091 AAGCCATGTGATCTCTACACAGTACTGTGAGCACATGCTGTGGTGAAGCTGGCTGTGG 69036
QY 600 AGACACAGGCTTAACCGTGTGATGGCTGACAGCTGCACCTGTTGTCATTTGGGTTGA 659
Db 69035 AGACACAGGCTTAACCGTGTGATGGCTGACAGCTGCACCTGTTGTCATTTGGGTTGA 68976
QY 660 CTTGTTTGGATTTGCTCTCTATGCTTAAATTTGCAAGCTGCTTTCCTTCCCTCTCATC 719
Db 68975 CTTGTTTGGATTTGCTCTCTATGCTTAAATTTGCAAGCTGCTTTCCTTCCCTCTCATC 68916
QY 720 CCATGAAGCTCGTCCAAAGCCCTAGGACCTGTGGTCCCATGCTGTGTGTCATCTCAT 779
Db 68915 CCATGAAGCTCGTCCAAAGCCCTAGGACCTGTGGTCCCATGCTGTGTGTCATCTCAT 68856
QY 780 CTCTTATACACAGCCCTCTTCTCTTTTACACAGCTTTGGCCATCACGTTCCAGT 839
Db 68855 CTCTTATACACAGCCCTCTTCTCTTTTACACAGCTTTGGCCATCACGTTCCAGT 68796
QY 840 CCATATTACATTTCTTTTGGCCAAATGTTTATCTCTTTTGGCCAAATGTTTATCTCTGT 899
Db 68795 CCATATTACATTTCTTTTGGCCAAATGTTTATCTCTTTTGGCCAAATGTTTATCTCTGT 68736
QY 900 GGTATATGGAGTTAAAGACCAACAGATCCGTAAAGAGTTGTCAAGGTGTTTCAAGTGG 959
Db 68735 GGTATATGGAGTTAAAGACCAACAGATCCGTAAAGAGTTGTCAAGGTGTTTCAAGTGG 68676
QY 960 GCAGGGAATGGGATCAAGCATCTGAG 987
Db 68675 GCAGGGAATGGGATCAAGCATCTGAG 68648
```

RESULT 14

LOCUS

AC147586

DEFINITION

Orolemur garnettii clone CH256-553F22, WORKING DRAFT SEQUENCE.

AC147586

ACCESSION

AC147586.1

VERSION

HTG; HTGS PHASE2; HTGS DRAFT.

KEYWORDS

Orolemur garnettii (small-eared galago)

SOURCE

Orolemur garnettii

ORGANISM

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Strepsirhini; Galagonidae; Orolemur.

REFERENCE

1 (bases 1 to 181386)

ANTHONY

Antony, A., Ayele, K., Benjamin, B., Blakesley, R. W.,

Bouffard, G. G., Brinkley, C., Brooks, S., Chu, G., Coleman, B.,

Coleman, H., Daki, N., Engle, J., Granite, S., Guan, X., Gupta, J.,

Haghighi, P., Han, J., Hansen, N., Ho, S.-L., Hu, P., Hurle, B.,

Idol, J. R., Jones, C., Karlins, E., Kim, H., Kwong, P., Latic, P.,

Larson, S., Lee-Lin, S.-Q., Legaspi, R., Maduro, Q. L., Maduro, V. B.,

Margulies, E. H., Masello, C., Maskeri, B., McDowell, J.,

Mullikin, J. C., Pagnirigan, C., Portnoy, M. E., Prasad, A., Puri, O.,

Maskeri, B., McDowell, J., Mullikin, J.C., Paguirigan, C., Pearson, R., Portnoy, M.E., Prasad, A., Reddix-Dugue, N., Schandler, K., Schueler, M.G., Shah, K., Sison, C., Stantripop, S., Thomas, J.W., Thomas, P.J., Tsipouri, V., Vogt, J.L., Wetherby, K.D., Young, A. and Green, E.D.

NISC Comparative Sequencing Initiative

Unpublished

2 (bases 1 to 239335)

Green, E.D.

Direct Submission

Submitted (11-SEP-2003) NIH Intramural Sequencing Center, 8717

Grovermont Circle, Gaithersburg, MD 20877, USA

3 (bases 1 to 239335)

Green, E.D.

Direct Submission

Submitted (17-OCT-2003) NIH Intramural Sequencing Center, 8717

Grovermont Circle, Gaithersburg, MD 20877, USA

On Oct 17, 2003 this sequence version replaced gi:34576313.

----- Genome Center

Center: NIH Intramural Sequencing Center

Center code: NISC

Web site: <http://www.nisc.nih.gov>Contact: nisc_zoo@nhgri.nih.gov

----- Project Information

Center project name: fgs

Center clone name: 105A13

The sequence data in this record represents an 'enhanced' version of a Phase 2 submission. Specifically, the indicated order and orientation of each sequence contig has been established using one or more of the following: read-pair data from individual subclones, overlaps with neighboring clones, alignment with available reference sequence (e.g., human), and/or confirmation by PCR testing. In addition, the sequence assembly is based on at least 8x average coverage in Q20 bases and has been reviewed to rule out gross misassemblies, the low-quality ends of sequence contigs have been trimmed away, and each base is associated with a Phrap-derived quality score.

----- Summary Statistics

Sequencing vector: plasmid; n/a; 100% of reads

Chemistry: Dye-terminator Big Dye; 100% of reads

Assembly program: Phrap; version 0.990319

Consensus quality: 238379 bases at least Q40

Consensus quality: 238694 bases at least Q30

Consensus quality: 238929 bases at least Q20

Insert size: 239000; agarose-fp

Insert size: 239035; sum-of-contigs

Quality coverage: 10.78x in Q20 bases; agarose-fp

Quality coverage: 10.78x in Q20 bases; sum-of-contigs

* NOTE: This is a 'working draft' sequence. It currently consists of 4 contigs. Gaps between the contigs are represented as runs of N. The order of the pieces is believed to be correct as given, however the sizes of the gaps between them are based on estimates that have been provided by the submitter.

* This sequence will be replaced

* by the finished sequence as soon as it is available and

* the accession number will be preserved.

* 1 4357: contig of 4357 bp in length

* 4358 4457: gap of unknown length

* 4458 6663: contig of 6206 bp in length

* 6664 66763: gap of unknown length

* 66764 232439: contig of 165676 bp in length

* 232440 232539: gap of unknown length

* 232540 239335: contig of 6796 bp in length.

Location/Qualifiers

1..239335

/organism="Otoleum garnettii"

/mol_type="genomic DNA"

/db_xref="taxon:30611"

/clone="CH256-105A13"

FEATURES

source

```
/clone_lib="CH256"
1..4357
/note="assembly_fragment"
clone_end:77
vector_side:left"
4458..6663
/note="assembly_fragment"
66764..232439
/note="assembly_fragment"
232540..239335
/note="assembly_fragment"
clone_end:SP6
vector_side:right"

ORIGIN

Query Match      83.3%; Score 822.2; DB 2; Length 239335;
Best Local Similarity 89.6%; Pred. No. 3.9e-238;
Matches 884; Conservative 0; Mismatches 103; Indels 0; Gaps 0;

QY 1 ATGTCCAGCACTCTTGGCCACACATGGAATCTCTCATCACATGATGTTGACCCCTTCT 60
Db 9997 ATATCCGGTACTCTTGGCCAGATCATGAAATCTCTTAATCACACACATGACCCCTTCT 9938
QY 61 GTCCTCTCTCTCTGGGCATCCAGGCTCGAACAAATTTCAITTTGGCTCTCATCCCT 120
Db 9937 ATTTCTTCTCTCTTGGGTATTCAGGCTCTAGAAAAATTTCAITTTGGCTCTCATCCCT 9878
QY 121 GTGTGTGGCTTTAGGCACAGCCACAAATTTGTGGCAATATACTATTCTCGTTGTTGTCGCC 180
Db 9877 GTGTGTGGCTTTAGGCACAGCCACAAATTTGTGGCAATATACTATTCTCGTTGTTGTCGCC 9818
QY 181 ACTGAACAGCTTTTGCACAGCCCTGTGTACCTTTTCTGTGCATGCTCTCAACCATCGAC 240
Db 9817 ACTGAGCCAGCCCTGTGTACCTTTTCTGTGCATGCTCTCAACCATCGAC 9758
QY 241 TTGCTGCTCTCTCTCCAGAGTCCCAAGTCTCTGCTATCTCTTGGTGTGGAGCCGGA 300
Db 9757 TTGCTGCTCTCTCTCCAGAGTCCCAAGTCTCTGCTATCTCTTGGTGTGGAGCCGGA 9698
QY 301 CATATATCTGCTCTGCTGCTGGCACATATGTTCTTTCATTCATGCTCTCGCATGATG 360
Db 9697 CATATATCTGCTCTGCTGCTGGCACATATGTTCTTTCATTCATGCTCTCGCATGATG 9638
QY 361 GAGTCCACTGTGTACTTGGCCATGCGCTTTGATCGTACGTGGCCATCTGCCACCCACTC 420
Db 9637 GAGTCCACTGTGTACTTGGCCATGCGCTTTGATCGTACGTGGCCATCTGCCACCCACTC 9578
QY 421 CGCTATGCCCAATCTCTCACTGACACCATCATTTGCCACATAGGGGTGGAGCTGTAGTG 480
Db 9577 CGCTATGCTACTATCTCTCACTGACACCATCATTTGCTGCGCATTTGAGTGGTATGGTG 9518
QY 481 CGAGGCTCCCTGCTCATGCTCCCATGTCCTTCTTTATTTGGGCGTTTGAATCTTGCCAA 540
Db 9517 CGAGGCTCCCTGCTCATGCTCCCATGTCCTTCTTTATTTGGGCGTTTGAATCTTGCCAA 9458
QY 541 AGCCATGTGATCTCTACACAGTACTGTGAGACATAGCTGTGTGAAGCTGGCTGTGGA 600
Db 9457 AGCCATGTGATCTCTACACAGTACTGTGAGACATAGCTGTGTGAAGCTGGCTGTGGA 9398
QY 601 GACACAGGCTTAACCGTGTGATGGGCTGACAGCTGACATGTTGGTTCATTTGGGGTGGAC 660
Db 9397 GACACAGGCTTAACCGTGTGATGGGCTGACAGCTGACATGTTGGTTCATTTGGGGTGGAC 9338
QY 661 TTGTTTGGCATTTGGTCTCTCTCTATGCGCTAAATTTGCAAGCTGTCCTTCGCTCTCATCC 720
Db 9337 TTGTTTGGCATTTGGTCTCTCTCTATGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 9278
QY 721 CATGAAGCTCGTCCAAAGGCCCTAGGACCTGTGTTTCCCATGCTGTGTGTCTCTCATC 780
Db 9277 CATGAAGCTCGTCCAAAGGCCCTAGGACCTGTGTTTCCCATGCTGTGTGTCTCTCATC 9218
QY 781 TCTTATACACAGCCCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 840
```

```

Db 9217 TCCTAGACACAGCTCTCTTTTCCTTTTACACATCGCTTTGGCCACCACGTTCCACTC 9158
Qy 841 CATATTCACATTCCTTTTGGCCAAATGTTTATCTGCTTTTGGCCACCTGCTCTTAAATCCTGTG 900
Db 9157 CATATTCACATTCCTTTTGGCCAAATGTTTATCTGCTTTTCCACCTGCTCTTAAACCTGTA 9098
Qy 901 GTATATCGAGTTAAGACCAACACAGATCCGTAAGAGAGTTGTCAGGGTGTTCAGAGTGGG 960
Db 9097 GTATATCGGGTTAAGACCAACAGAGATTCGTGAAGGGTTGTCAGAGTGTTCAGAGAGGA 9038
Qy 961 CAGGAATGGGCATCAAGGCATCTGAG 987
Db 9037 CAGGAATTGGAATCAGGCATCTGAG 9011

```

Search completed: August 27, 2004, 17:37:04
 Job time : 4178 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 27, 2004, 16:10:03 ; Search time 3091 Seconds

(without alignments)

9535.416 Million cell updates/sec

Title: US-10-081-775-1_COPY_537_1523

Perfect score: 987

Sequence: 1 atgtccagactcttggcca.....tgggcataaggaatcttgag 987

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 27513289 seqs, 14931090276 residues

Total number of hits satisfying chosen parameters: 55026578

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

EST:*

1: em_estba:*

2: em_esthum:*

3: em_estin:*

4: em_estmu:*

5: em_estov:*

6: em_estpl:*

7: em_estro:*

8: em_estc:*

9: gb_est1:*

10: gb_est2:*

11: gb_est3:*

12: gb_est4:*

13: gb_est5:*

14: gb_est6:*

15: em_estfun:*

16: em_estom:*

17: em_gss_hum:*

18: em_gss_inv:*

19: em_gss_pln:*

20: em_gss_vrt:*

21: em_gss_fun:*

22: em_gss_mam:*

23: em_gss_mus:*

24: em_gss_pro:*

25: em_gss_rod:*

26: em_gss_phg:*

27: em_gss_vrl:*

28: gb_gss1:*

29: gb_gss2:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	346.6	35.1	948	29 AY401474	AY401474 Mus muscu
2	345	35.0	948	29 AY401472	AY401472 Homo sapi
3	344	34.9	948	29 AY410600	AY410600 Mus muscu
4	340.4	34.5	948	29 AY410598	AY410598 Homo sapi

5	337.2	34.2	948	29 AY410599	AY410599 Pan trogl
6	336.4	34.1	948	29 AY401473	AY401473 Pan trogl
7	331.2	33.6	957	29 AY414225	AY414225 Homo sapi
8	327	33.1	1964	11 AK036356	AK036356 Mus muscu
9	327	33.1	3410	11 AK028467	AK028467 Mus muscu
10	314	31.8	822	29 CC500683	CC500683 CH240_338
11	310	31.4	936	29 AY401478	AY401478 Homo sapi
12	307.8	31.2	954	29 AY414227	AY414227 Mus muscu
13	300.6	30.5	936	29 AY401480	AY401480 Mus muscu
14	297.8	30.2	774	29 CC512357	CC512357 CH240_356
15	294.4	29.8	900	29 AY410595	AY410595 Homo sapi
16	293.6	29.7	966	29 AY405569	AY405569 Mus muscu
17	291.4	29.5	900	29 AY410596	AY410596 Pan trogl
18	284.2	28.8	846	29 AY402272	AY402272 Mus muscu
19	283.6	28.7	966	29 AY405567	AY405567 Homo sapi
20	283	28.7	900	29 AY410597	AY410597 Mus muscu
21	281.6	28.5	846	29 AY402270	AY402270 Homo sapi
22	272.6	27.6	957	29 AY414226	AY414226 Pan trogl
23	270.2	27.4	1086	28 AF101706	AF101706 AF101706
24	256.4	26.0	746	29 AY401479	AY401479 Pan trogl
25	254.6	25.8	758	29 CE341239	CE341239 tigr-gss-
26	251.2	25.5	864	29 CC480759	CC480759 CH240_308
27	238.6	24.2	836	29 CC526526	CC526526 CH240_401
28	236	23.9	798	28 BH082934	BH082934 RPCI-24-9
29	231.4	23.4	939	29 AY402268	AY402268 Pan trogl
30	229.8	23.3	939	29 AY402267	AY402267 Homo sapi
31	229.6	23.2	948	29 AY402752	AY402752 Mus muscu
32	228.8	23.2	820	28 BH046083	BH046083 RPCI-24-2
33	227.8	23.1	966	29 AY401649	AY401649 Homo sapi
34	218.8	22.2	853	29 CC527411	CC527411 CH240_402
35	217.2	22.0	963	29 AY401651	AY401651 Mus muscu
36	216.6	21.9	825	28 AF156720	AF156720 AF156720
37	213.2	21.6	948	29 AY402750	AY402750 Homo sapi
38	213	21.6	701	29 CE070260	CE070260 tigr-gss-
39	211.8	21.5	942	29 AY402281	AY402281 Mus muscu
40	211.4	21.4	709	29 CC503779	CC503779 CH240_343
41	211	21.4	669	29 CC569584	CC569584 CH240_444
42	209.2	21.2	705	28 AF101881	AF101881 AF101881
43	207.8	21.1	786	29 AY414228	AY414228 Homo sapi
44	206.6	20.9	822	29 AY405759	AY405759 Homo sapi
45	203.6	20.6	795	28 BZ256668	BZ256668 CH230-334

ALIGNMENTS

RESULT 1	AY401474	948 bp	DNA	linear	GSS 12-DEC-2003
LOCUS	AY401474				
DEFINITION	Mus musculus HCM0902 gene, VIRTUAL TRANSCRIPT, partial sequence, genomic survey sequence.				
ACCESSION	AY401474				
VERSION	AY401474.1	GI:39757463			
KEYWORDS	GSS.				
SOURCE	Mus musculus (house mouse)				
ORGANISM	Mus musculus				
REFERENCE	Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.				
AUTHORS	1 (bases 1 to 948) Clark, A.G., Gnanowski, S., Nielson, R., Thomas, P., Kejarival, A., Todd, M.A., Tanenbaum, D.M., Civello, D.R., Lu, F., Murphy, B., Ferriera, S., Wang, G., Zheng, X.H., White, T.J., Sninsky, J.J., Adams, M.D. and Cargill, M. Inferring nonneutral evolution from human-chimp-mouse orthologous gene trios				
TITLE	Science 302 (5652), 1960-1963 (2003)				
JOURNAL	14671302				
PUBMED	2 (bases 1 to 948)				
REFERENCE	Clark, A.G., Gnanowski, S., Nielson, R., Thomas, P., Kejarival, A., Todd, M.A., Tanenbaum, D.M., Civello, D.R., Lu, F., Murphy, B., Ferriera, S., Wang, G., Zheng, X.H., White, T.J., Sninsky, J.J., Adams, M.D. and Cargill, M. Direct Submission				
AUTHORS					
TITLE					

JOURNAL	Submitted (16-NOV-2003) Celera Genomics, 45 West Gude Drive, Rockville, MD 20850, USA						
COMMENT	This sequence was made by sequencing genomic exons and ordering them based on alignment.						
FEATURES	Location/Qualifiers						
source	1..948 /organism="Mus musculus" /mol_type="genomic DNA" /db_xref="taxon:10090" <1..>948 /locus_tag="HMC0902"						
gene							
ORIGIN							
Query Match	35.1%; Score 346.6; DB 29; Length 948;						
Best Local Similarity	61.4%; Pred. No. 3.6e-86;						
Matches	575; Conservative 0; Mismatches 359; Indels 3; Gaps 1.						
Qy	25	ATGGAATCTCTCATCACACTGATGTTCAGCCCTTCTGCTTTCTTCTCCTCGGGATCCCA	84				
Db	1	ATGCAACACAAACCAGGCATCAGAACCAGCTCTCTCTGCTCATGGGATTCCC	60				
Qy	85	GGTCTGAAACAATTTCATTTTGCGTCTCACTCCCTGTGTGTGGCTTAGGCACACGCCA	144				
Db	61	GGCCTGGAGGCGTCCCACCTTTTGGATTGCAATTCCTTCTGTTCATGTATGCCCCAGCA	120				
Qy	145	ATTGTGGCAATAFAACTATTTCTGGTGTGTTGTGGCCACTGAACCAAGTCTTGCACAAGCCT	204				
Db	121	GTTCTTGGCAACATGGCAGTCTGTGTGTGTGGCTCTGAGCCTTCACTGCACCAACC	180				
Qy	205	GTGTACCTTTTTCTGTGCATGCTCTCAACCACTCACTTGGCTGGCTCTGTCTCCACAGTT	264				
Db	181	ATGTACCTGTTCCTGTGTGATGCTGTCCAATTATGACCTGTAFACTCTGCACTTCCACTGTC	240				
Qy	265	CCCAAGCTACTGGGTATCTTCTGTGTGGAGCGGACATATATCTGCCCTCTGCTGCGCTG	324				
Db	241	CCCAAGCTCCTCGCACCTTTTGGCAAATGCAGCTGAGATTGCCCTTTGGAGCCTGTGCT	300				
Qy	325	GCACATATGTTCTTCATTCATGCTTCTGCAATGATGGAGTCCACTGTGTGCTACTGCGCAATG	384				
Db	301	ACCCAGATGTTCTTTATCCATGGTTCTTCGGCTGTAGAATCCGGTATCCGTGATCAATG	360				
Qy	385	GCCTTTGATGCTACGTGGCCATCTGCCACCCACTCCGCTATGCCACAATCTCACTGCAC	444				
Db	361	GCCTTTGATGCTACTTGGCCATCTGCCGACCAATTCGACTACGGGTCAATTGCTGTCTCA	420				
Qy	445	ACCATCATTTGCCACATAGGGGTGGCAGCTGTAGTGGAGGCTCCCTGCTCATGTGCCCA	504				
Db	421	GAGTCTGTGACAAAGTTAGGAGCTGCTGCTCTGCTCCGTGGCTCGGGCTCATGACCCCA	480				
Qy	505	TGTCCTTCTTTATTGGCGGTTTGAACCTTCTGCCAAAGCCATGTGATCCTACACAGGTAC	564				
Db	481	CTTACTCGCTACTGGCAAGGCTGAGCTACTG---TGCGCGAGTGGTAGCAATTTCTCTAC	537				
Qy	565	TGTGAGCATGCGTGTGCTGAAGCTGCGCTGTGGAGACACCAGGCCTTAACCGTGTGTAT	624				
Db	538	TGTGAGCATGCGTGTGCTGAAGCTGCGCTGTGGGGGACACCAACCAACACATATAT	597				
Qy	625	GGGCTGACAGCTGCACTGTTGGTCAATGGGGTTGACTTTTGTGCAATGGTCTCTCTAT	684				
Db	598	GGCATCACAGCAGCCACACTGGTGGTAGGAAGTGAATCCATTTGTATTGGCCATATCTCTAT	657				
Qy	685	GCCCTAATTGCACAAGCTGTCTTCGCTCTCATCCCATGAGCTCGGTCCAGGCCCTTA	744				
Db	658	GCATCATCTCTCCGGGCTGTGTGTGGGCTCTCTCTCCAAAGAGGCCAGGGCAAGACCTTT	717				
Qy	745	GGGACCTGTGGTTCCTCATGTCTGTGCTATCCCTCATCTCTTATACACCAAGCCCTCTTCTCC	804				
Db	718	GGCATGTGTGGCTCCCACTGGGTGTATCTCTCTCTCTATACACAGGGCTCTTCTCA	777				
Qy	805	TTTTTTACACACCGCTTTGGGCAATCAAGTTTCAGTCCATATTTCAATTTCTTTTGGCCAAT	864				
Db	778	TTCTATACCCAGCGTTTGGTCAACATGTACCCAGGACGTCACATCTCTCTGGTGTGAC	837				


```
QY 232 ACCATCGACTGGCTGCTCTGTCTCCACAGTTCCTCCAGCTACTGGCTATCTCTGGTGT 291
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
121 CTCAGGACCTGGGATGTCACACACACATCCCAAGGTGTGTGATCTTCTGGTGT 180
QY 292 GGAGCGGACATATATCTGCTCTGCTGCTGCTGGCAGATATGTTCTTCAATTCATGCTTC 351
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
181 GACCAGAGTGACATCAGCTTCGAGGGCTGCTGGTCCAGCTGTCTTCTCCACTCCATC 240
QY 352 TGCATGATGAGTCCACTGTGCTACTGCGCATGGCTTTGATCGTACGTGGCCATCTGC 411
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
241 TCTGCTTGCAGTCTGCCATCTCATGTCCTATGCTGCTGCTGCTGCTGCTGCTGCTGCT 300
QY 412 CACCCACTCCGCTATGCCAATCTCTACTGACACCATCATTCGCCACATAGGGGTGCA 471
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
301 GAGCCCTGCGCTATGCCACCATCTTCCACAGCCGCTATGGCTCATCGGCTCTGTG 360
QY 472 GCTGTAGTGGAGGCTCCCTGCTCATGCTCCCATGCTGCTCTTCTTATTTGGGCTTTGAAC 531
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
361 AGTTTAGTGGAGCTGCTGCTCATCTCCCATGCCATCTCTCCACAGATGCC 420
QY 532 TTCTGCCAAGCATGATCTTACACAGTACTGTGAGACATGCGTGTGGTGAAGCTG 591
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
421 TTTCATGCCAGCATGTATCCCAACCACTGCTGTGAGCAGTGGCTGTGGTGAAGATG 480
QY 592 GCCTGTGGAGACACAGGCTTAACCGTGTATGGCTGACAGCTGACACTGTGTGCTATT 651
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
481 GTGTGTGGAGACACAGGCTCAACAGGTATGCTGTGGTGGGCTTGTGTGCT 540
QY 652 GGGGTGACTGTTGTTGCAITGGTCTCTCTATGCTGCTTAAATGCAACAGCTGTCTTCGC 711
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
541 GGGGTAGACACTCAGCTATTGGCTCATCTTATGTGCTGATCATCCGGGCTGTAATGCG 600
QY 712 CTCTATCCCATGAGCTGGTCCAGGCTTAGGACCTGTGGTCCCATGCTGTGCTGTC 771
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
601 CTCTCTTCTAAGGAGCCCATCAAAAGAGCAATAATACCTGCACACACATCTGTGTC 660
QY 772 ATCTCTATCTTATACACAGGCTCTCTCTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 831
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
661 ATGCTTATCTTATCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 720
QY 832 GTTCCAGTCCATATTCATCTTCTTCTGCTGCTTCTTCTGCTTCTTCTGCTTCTTCTGCT 891
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
721 ATCCACCTCATGTGACACCATCTTGGCAGCTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 780
QY 892 AATCTGTGATATGAGTTAAGACCAACAGATCCG 929
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
781 AACCCATTAATTTATGCAAGTAAACCAAGAGTTTCG 818
```

```
RESULT 11
LOCUS AY401478
DEFINITION Homo sapiens HCM0904 gene, VIRTUAL TRANSCRIPT, partial sequence,
ACCESSION AY401478
VERSION AY401478.1 GI:39757467
KEYWORDS GSS.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 936)
Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejeriwal,A.,
Todd,M.A., Tanenbaum,D.M., Civello,D.R., Lu,F., Murphy,B.,
Ferriera,S., Wang,G., Zheng,X.H., White,T.J., Sninsky,J.J.,
Adams,M.D. and Cargill,M.
Inferring nonneutral evolution from human-chimp-mouse orthologous
gene trios
JOURNAL Science 302 (5652), 1960-1963 (2003)
PUBMED 14671302
REFERENCE 2 (bases 1 to 936)
AUTHORS Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejeriwal,A.,
```

```
Todd,M.A., Tanenbaum,D.M., Civello,D.R., Lu,F., Murphy,B.,
Ferriera,S., Wang,G., Zheng,X.H., White,T.J., Sninsky,J.J.,
Adams,M.D. and Cargill,M.
Direct Submission
Submitted (16-NOV-2003) Celera Genomics, 45 West Gude Drive,
Rockville, MD 20850, USA
This sequence was made by sequencing genomic exons and ordering
them based on alignment.
FEATURES
source
1..936
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
<1_>936
/locus_tag="HCM0904"
Query Match 31.4%; Score 310; DB 29; Length 936;
Best Local Similarity 59.8%; Pred. No. 7.5e-76;
Matches 539; Conservative 0; Mismatches 360; Indels 3; Gaps 1;
```

```
QY 50 TTGACCCCTTCTGCTCTTCTTCTCTGCGCATCCAGGCTCGAACAATTTCAATTTGGC 109
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
26 TTCACCCAGTCACATTTTCTCTCATTTGGAATCCAGGCTCGGAAGACTTCCACATGTGGA 85
QY 110 TCTCACTCCCTGCTGTGGCTTAGGCACAGCCAAATTTGGCAATATACTATTCTGG 169
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
86 TCTCGGGCTTTCTGCTCTGCTGCTTACCTTGTGGGTTTGTGGGCAATGCCACCATCTGC 145
QY 170 TTGTTGTTGGCACTGAACCACTTTGCAACAGCTGTGTACTCTTTTCTGTGCAATGCTCT 229
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
146 TAGTCATCAAGTAGAAGACAGACTCTCGGGAGCCATGTTCTACTTCTCGGCCATCTTT 205
QY 230 CAACCATCGACTTGGCTGCTCTGCTCCACAGTTCCCAAGCTACTGCTATCTTCTGGT 289
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
206 CCACATTTGATTTGGCCCTTTCTGCAACCTCTGTGCTCGCATGCTGGGTATCTTCTGGT 265
QY 290 GTGGAGCGGACATATATCTGCTCTGCTCGCTGGCACATATGTTCTTCAATCATGCTCT 349
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
266 TTGATGCTCAGAGATTAACTATGAGGCTGTGTGGCCAGATGTTTCTGATCATGCTCT 325
QY 350 TCTGCATGATGAGTCCACTGTGCTACTGGCCATGGCTTTGATCGCTACGTGCCATCT 409
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
326 TCACTGGCATGGAGCTGAGGTCTTACTGGCTATGGCTTTGACCGTTATGTGGCATCT 385
QY 410 GCCACCACTCCGCTATGCCACATCTCTACTGACACCATCATTTGCCACAGGGGTGG 469
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
386 GTGCTCCACTACATTTACCAACCATCTTGACATCCCTAGTGTGGTGGGCATTTAGCATGT 445
QY 470 CAGCTGTAGTCGAGGCTCCCTGCTCATGCTCCCATGCTCTTCTTATTTGGGCTTTGA 529
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
446 GCATTGTAAATCGTCCGCTTTTACTTACATTTCCATGGTCTATCTTATCTACGCTTAC 505
QY 530 ACTTCTGCCAAGCCATGTGATCTCTACACAGTACTGTGAGCAGCATGCTGTGGTGAAGC 589
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
506 CTTTTGTGAGGCTCACATAATAGCCCATCTCTACTGTGAGCAGCATGGCATTTGCAAAAT 565
QY 590 TGGCTGTGGAGACACAGGCTAACCGTGTGTATGGCTGACAGCTGCATGTGTGCTCA 649
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
566 TGTCTCTGGAAACATTCGTATCAATGGTATCTATGGGCTTTT---TGTAGTTCTTTCT 622
QY 650 TTGGGGTGTACTTGTGTTGCTCTCTCTATGCTTAAATTTGCAACAGCTGTCTCTTC 709
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
623 TTGTTCTGAACCTGTGCTCATTTGGCATCTCGTATGTTTACATTTCCGTCGTCTCTCC 682
QY 710 GCCTCTCATCCATGAGCTGGTCCAGGCTCTAGGACCTGTGGTTCCTCATGTCTGTG 769
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
683 GCCTCCCATCATGATGCTCAGCTAAAAGCCCTTAAGCAGCTGTGGGCTCATGTGTGAG 742
QY 770 TCATCCTCATCTTATACACAGCCCTCTTCTCTCTTTTACACACCGCTTTTGCCATC 829
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
743 TCATCTGTGTTTCTATATCCCTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 802
```

```

QY 830 ACGTTCAGTCCATATTCACATCTTTTGGCAATGTTATCTGCTTTTGGCACTGCTC 889
Db 803 AAATACCAGGTTACATTCACATCTTTGTTGCAATCTCTATTTGATTATCCACCCCTCTC 862
QY 890 TTAATCCTGTGTATATGAGTAAAGACCAACAGATCCGTAAAGAGTTGTCAAGGTTGT 949
Db 863 TCAACCCCATATTATGAGGTGAGGACCAACAGATTCGAGAGCGAGTGTCTATGTTT 922
QY 950 TT 951
Db 923 TT 924

RESULT 12
LOCUS AY414227 954 bp DNA linear GSS 12-DEC-2003
DEFINITION Mus musculus HCM5153 gene, VIRTUAL TRANSCRIPT, partial sequence,
genomic survey sequence.
ACCESSION AY414227
VERSION AY414227.1 GI:39770189
KEYWORDS GSS.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 954)
Clark, A.G., Glanowski, S., Nielson, R., Thomas, P., Kejarial, A.,
Todd, M.A., Tanenbaum, D.M., Civello, D.R., Lu, F., Murphy, B.,
Ferriera, S., Wang, G., Zheng, X.H., White, T.J., Sninsky, J.J.,
Adams, M.D. and Cargill, M.
Inferring nonneutral evolution from human-chimp-mouse orthologous
gene trios
JOURNAL Science 302 (5652), 1960-1963 (2003)
PUBMED 14671302
REFERENCE 2 (bases 1 to 954)
AUTHORS Clark, A.G., Glanowski, S., Nielson, R., Thomas, P., Kejarial, A.,
Todd, M.A., Tanenbaum, D.M., Civello, D.R., Lu, F., Murphy, B.,
Ferriera, S., Wang, G., Zheng, X.H., White, T.J., Sninsky, J.J.,
Adams, M.D. and Cargill, M.
TITLE Direct Submission
JOURNAL Submitted (16-NOV-2003) Celera Genomics, 45 West Gude Drive,
Rockville, MD 20850, USA
COMMENT This sequence was made by sequencing genomic exons and ordering
them based on alignment.
FEATURES
source Location/Qualifiers
1..954
/organism="Mus musculus"
/mol_type="genomic DNA"
/db_xref="taxon:10090"
<1..>954
/locus_tag="HCM5153"

gene
ORIGIN
Query Match 31.2%; Score 307.8; DB 29; Length 954;
Best Local Similarity 59.6%; Pred. No. 3.1e-75;
Matches 538; Conservative 0; Mismatches 362; Indels 3; Gaps 1;

QY 65 TCTTCTCCTGGGATCCAGTCTGGAAACAATTTCATTTGGGCTCTCATCCCTGTGT 124
Db 41 TCTTCTCAACTGGGATCCAGGCTGGAGTTTCGCCCACTCTGGATTGCCATCCCTTCT 100
QY 125 GTGGCTTAGGCACACCCACCAATTTGGGCAATATTAATCTTCTGTTGTTGGCACTG 184
Db 101 GTGTCAATGATGTGGTCTCTGCTTGGGAATGCTGCGCCTCATCTTATCATTTGGGAG 160
QY 185 AACCACTCTGCACAAAGCTGTGTAACCTTTTCTGTGCATGCTCTCAACCACTGACTGG 244
Db 161 AGAGTGTCTTACACACCCATGACCTTCTCTCTGCTTCTCTCTCATCTGACTGTGG 220
QY 245 CTGCTCTGTCTCCACAGTCTCCCAAGCTACTGGCTATCTTCTGTTGGAGCGGACATA 304
Db 221 CTCTCAGCTCCACCACTGTACCTAAAGTTGGCCATCTTATGGCTCCATCTTAATGAGA 280

```

```

QY 305 TATCTGCTCTGCTGCTGCGCACATATGTTCTTCAATTCATGCTTCTGCAATGAGT 364
Db 281 TTTCTTGTGGTGGATGCTGCTGCGCAGATGTTTGTGCCATCTCCATCTATGCTCTGAGT 340
QY 365 CCAGTGTCTACTGGCCATGGCTTTGATCGCTACGTGGCCATCTGCCACCCACCTCCGCT 424
Db 341 CCTCGTCTCTTGGCCATGGCTTTGATCGATATGATAGCTATCTGCAACCCACTGAGAT 400
QY 425 ATGCCAATCCTCACTGACACCAATCAATGCCCCACATAGGGGTGGAGCTGTAGTGGAG 484
Db 401 ATACAAACCATCTCAACCATACTGTCTATAGCCCAAAATATCTTTGCTGGGATAGTCCGTA 460
QY 485 GCTCCCTGCTCATGCTCCCATGCTCCCTTTTATTTGGGCGTTTGAATCTTCCGCAAGCC 544
Db 461 GTGGGCTATTGTTCTCCCTTTTCATCTTTTGTAGCGGACTGCCCTACTGTGTGTCACC 520
QY 545 ATGTGATCTTACACACGCTACTGTGAGCACATGGCTGTGGTGAAGCTGGCCCTGTGAGACA 604
Db 521 GTGTCACTGACTCACACATCTGTGAGCATATGGCATTTGCTAGGCTTGCCTTGCCCAATA 580
QY 605 CCAGGCTAACCGTGTGTATGGGCTGACAGCTGCACTGTGTGCTCATTTGGGGTTGACTTGT 664
Db 581 TTACTGTAAATATTGTTCTATGGGTTGACTGTAGCTTTGCTGGCCATGGGCTGGATTCCA 640
QY 665 TTTGCATTGGTCTCTCCTATGCCCTAATTGACAAAGCTGCTTCGCTCTCATCCCATG 724
Db 641 TCCTCATTTGCCATCTCCTATGGCTTCATCTCCGGGCTGCTTCGCTTCCATCCCGTG 700
QY 725 AAGCTTCGCTCAAGGCCCTTAGGGACCTGTGGTTCCTCATGCTGTGTCTCATCTCTCTT 784
Db 701 ATGCTCAACACAAAGCTCTGAGTACCTGTGGATCTCACTTGGTGTCTATCTGGTTTCT 760
QY 785 ATACACAGGCTCTCTCTCTCTTTTATACACAGCTTTGG---CCATCAGTTCAGTCC 841
Db 761 ATATACCTGCTTTCTCTCTCTCCCTACCCATCGCTTTGGTCACAATCGAGTTCCTCAAGC 820
QY 842 ATATTCACTTTTGGCCAACTTTATCTGCTTTTGGCCACCTGCTCTTAACTCTGTTGG 901
Db 821 ATGTGCATATCTCTCGCTAACCTCTATGTGTGGTGGCTCCCGTGTCTCAATCCCACTCA 880
QY 902 TATATGAGTGTAAAGACAAACAGATCCGTAAAGAGTTGTTCAGGGTGTTCAAAAGTGGGC 961
Db 881 TCTAGGTGCAAGAACCAAGAGATTCGAGTCGGCTCTTAAACTGCTTCACTTGGGGA 940
QY 962 AGG 964
Db 941 AGG 943

RESULT 13
LOCUS AY401480 936 bp DNA linear GSS 12-DEC-2003
DEFINITION Mus musculus HCM0904 gene, VIRTUAL TRANSCRIPT, partial sequence,
genomic survey sequence.
ACCESSION AY401480
VERSION AY401480.1 GI:39757469
KEYWORDS GSS.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 936)
Clark, A.G., Glanowski, S., Nielson, R., Thomas, P., Kejarial, A.,
Todd, M.A., Tanenbaum, D.M., Civello, D.R., Lu, F., Murphy, B.,
Ferriera, S., Wang, G., Zheng, X.H., White, T.J., Sninsky, J.J.,
Adams, M.D. and Cargill, M.
Inferring nonneutral evolution from human-chimp-mouse orthologous
gene trios
JOURNAL Science 302 (5652), 1960-1963 (2003)
PUBMED 14671302
REFERENCE 2 (bases 1 to 936)
AUTHORS Clark, A.G., Glanowski, S., Nielson, R., Thomas, P., Kejarial, A.,
Todd, M.A., Tanenbaum, D.M., Civello, D.R., Lu, F., Murphy, B.,

```

Ferrera,S., Wang,G., Zheng,X.H., White,T.J., Sninsky,J.J.,
 Adams,M.D. and Cargill,M.
 Direct Submission
 Submitted (16-NOV-2003) Celera Genomics, 45 West Gude Drive,
 Rockville, MD 20850, USA
 This sequence was made by sequencing genomic exons and ordering
 them based on alignment.
FEATURES
 source
 location/Qualifiers
 1..936
 /organism="Mus musculus"
 /mol_type="genomic DNA"
 /db_xref="taxon:10090"
 <!-->936
 /locus_tag="HCM0904"
gene
ORIGIN

 Query Match 30.5%; Score 300.6; DB 29; Length 936;
 Best Local Similarity 59.0%; Pred. No. 3.3e-73;
 Matches 535; Conservative 0; Mismatches 369; Indels 3; Gaps 1;

 QY 36 TCATCACATGATGTACCGCTTCTGCTCTTCTCCTGGGCATCCAGGCTCTGAACA 95
 ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
 Db 12 TAATAGAGCGTTTCTCACCAGATACATTTTCTCATGGGATCCAGGCTCTGGAGA 71

 QY 96 ATTTCATTGTGGCTCTCACTCCCCTGTGTGGCTTAGGCACAGCACAAATGTGGGCAA 155
 ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
 Db 72 AATCCATGCCGTGGATTCTCGCCTTCTCGCTGATTATCTTGCGCCCTAATGGTAAA 131

 QY 156 TATAACTATTCTGGTTGTTGGCACTGAACCAAGCTWTGGCACAAAGCCTGTGTACCTTTT 215
 ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
 Db 132 TACCATGATCTCGTAGTATTACAAGACAGAGCGTTCTCTCGAGAGCCCATGTCTACTT 191

 QY 216 TCTGTGATGTCCTCAACCATCGACTGGCGCTCTGTCTCCACAGTTCCCAAGCTACT 275
 ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
 Db 192 TCTGGCCATCCTTCTTCAGTGCACTTGGCTCTTTTCTAACACCTCTGTGCCCCGTATGCT 251

 QY 276 GGCTATCTTCTGGTGTGGAGCGGACATATATCTGCCTCTGCCTCGCTGGCACATATGTT 335
 ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
 Db 252 TGGCATATTCTGGTTTGATGCCCATGAATAAACTTGGTGGCTGTGTGGCTCGAGATGTT 311

 QY 336 CTTCATTTCATGCCCTTCTGCATGATGGAGTCCACTGTGCTACTGGCCATGGCCTTTGATCG 395
 ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
 Db 312 TTTTGATCCATGCCCTTACTGGCATGGAGGCTGAGGTTCTCATGGCCATGGCTTTTGACCG 371

 QY 396 CTACGTGGCCATCTGCCACCCACCTCCGCTATGCCACAATCCTCACTGACACCATCATGTC 455
 ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
 Db 372 TTATGTAGCCATCTGTGCTCCACTCCACTACACACCATCTTGACATCTCGTGTCTGAT 431

 QY 456 CCACATAGGGGTGGCAGCTGTAGTGCAGAGGCTCCCTGCTANGCTCCCAATGCCCTGCTT 515
 ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
 Db 432 GGGCATCAGCATATGTGTGTAATTCGTCCAGCCCTGTGTTATATGCCCAATGGTCTATCT 491

 QY 516 TATTGGCGGTTTGAACTTCTGCCAAAGCATGTGATCTACACACGTA CTGTGAGCACAT 575
 ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
 Db 492 CATTTACCGCTTACCTTTTGTCAAGCTCATGTAA TAGCCCATTCCTATTGTGAAACAT 551

 QY 576 GGCTGTGGTGAAGCTGGCCCTGTGAGACACACAGGCCCTAACCGTGTATGGGCTGACAGC 635
 ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
 Db 552 GGGCAITGCAAAATGTCTGTGAGACATCCATATCAACCGCTGTATGGCCT---TTT 608

 QY 636 TGCACTGTGTGTCAATTGGGTTGACTGTTTGTGATTGGTCTCTCCCTATGCCCCTAATGTC 695
 ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
 Db 609 TGTGGTCTCTCTTTCTTTTGAACCTTGGTACTTATTGGCACTCTCATATGGGTACATTTCT 668

 QY 696 ACAAGCTGTCTTCGCCCTCTCATCCCCATGAAGCTCGGTCCAAGGCCCTTAGGAGCCTGTGG 755
 ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
 Db 669 TCGTGCTGTGTTTTCGCTCCCATCTCAGGATGCACGGTTAAAAAGCCCTCAGCACATGTGG 728

 QY 756 TTCCCATGTCTGTGTCACTCTCTTATACAGAGCCCTCTTCTCCTTTTTTTACACA 815
 ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
 Db 729 TTCCCATGTGTCTGTCTGTGTGTTTTCTACATCCCATCACTGCTCTCTTTCTTCCCTCACTCA 788

 QY 816 CCGCTTTGGCCATCAGCTTCCAGTCCCATTTACATTTCTTTTGGCCAAATGTTTATCTGCT 875

Db 721 GGTACCTGTGCTCTCACATCAGTGCAGTTGCCAICTTCTACATCCCATGATCAGTTG 780
Qy 805 TTTTTCACACACCGCTTTGGCCATCAGTTCAGTTCATATTCACATTCCTTTGGCCAAT 864
Db 781 TCACTAGTGCACAGATTTGGAAAGCATGCCCTCCCTTTGTGCACATGCTCATTGCCAAT 840
Qy 865 GTTTATCTGCTTTTGGCCACCTGCTCTTAATCCTGTGGTATATGAGTTAAGACCAA 920
Db 841 GTTTATCTGCTTATCCCTCCTGTAAATGAATCCCATATCTACAGTGTAAAGACCAA 896

Search completed: August 27, 2004, 18:28:39
Job time : 3097 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: August 27, 2004, 17:37:09 ; Search time 4663 Seconds
(without alignments)
3058.087 Million cell updates/sec

Title: US-10-081-775-2

Perfect score: 1718

Sequence: 1 MSSTLGHNNESPHHTDVPDPS.....RKRVRVFGSGMGKASE 329

Scoring table:

BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 3470272 seqs, 21671516995 residues

Total number of hits satisfying chosen parameters: 6940544

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+p2n.model -DEVEL=xlh
-Q=/cgn2_1/USPTO.spool/US10081775/runat_20082004_170213_11126/app query.fasta_1.519
-DB=GenEmbl -QMT=fastap -SUFFIX=rge -MINMATCH=0.1 -LOOPEL=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=-1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45
-DOCALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MTN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US10081775@cgn_1_1_3731@runat_20082004_170213_11126 -NCPU=6 -ICPU=3
-NO_MMAP -LARGQUERY -NEG SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

GenEmbl:

1: gb.ba.*
2: gb.htg.*
3: gb.in.*
4: gb.om.*
5: gb.ov.*
6: gb.pat.*
7: gb.ph.*
8: gb.pl.*
9: gb.pr.*
10: gb.ro.*
11: gb.sts.*
12: gb.sy.*
13: gb.un.*
14: gb.vi.*
15: em.ba.*
16: em.fun.*
17: em.hum.*
18: em.in.*
19: em.mu.*
20: em.om.*
21: em.or.*
22: em.ov.*
23: em.pat.*
24: em.ph.*
25: em.pl.*
26: em.ro.*
27: em.sts.*
28: em.un.*

29: em.vi.*
30: em.htg.hum.*
31: em.htg.inv.*
32: em.htg.other.*
33: em.htg.mus.*
34: em.htg.pln.*
35: em.htg.rat.*
36: em.htg.rod.*
37: em.htg.mam.*
38: em.htg.vrt.*
39: em.sy.*
39: em.htgo.hum.*
40: em.htgo.mus.*
41: em.htgo.other.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1705	99.2	185330	2	AC124272 Homo sapi
2	1705	98.2	203352	2	AC111179 Homo sapi
3	1691	98.4	990	6	BD144452 Novel G-p
4	1691	98.4	998	6	AX551417 Sequence
5	1691	98.4	998	6	AX551419 Sequence
6	1691	98.4	1390	6	AX646605 Sequence
7	1691	98.4	1390	6	AB065534 Homo sapi
8	1691	98.4	172027	2	AC026090 Homo sapi
9	1691	98.4	178597	9	AC116156 Homo sapi
10	1649	96.0	963	6	AX241675 Sequence
11	1649	96.0	966	6	AX318261 Sequence
12	1649	96.0	966	6	AX448757 Sequence
13	1647.5	95.9	202761	9	AC109341 Homo sapi
14	1558	90.7	181386	2	AC147586 Otolenur
15	1558	90.7	239335	2	AC146635 Sequence
16	1534	89.3	990	10	AY073012 Mus muscu
17	1534	89.3	191432	2	AC135109 Mus muscu
18	1532	89.2	264026	2	AC113720 Rattus no
19	1517	88.3	966	10	AY317777 Mus muscu
20	1376	80.1	948	6	AX659373 Sequence
21	1224	71.2	781	4	AY355971 Canis fam
22	979.5	57.0	227938	2	AC098008 Rattus no
23	978.5	57.0	1300	6	AX646641 Sequence
24	978.5	57.0	1300	9	AB065856 Homo sapi
25	978.5	57.0	141485	9	AC009758 Homo sapi
26	977.5	56.9	948	10	AY073340 Mus muscu
27	977.5	56.9	948	10	AY317664 Mus muscu
28	977.5	56.9	256243	2	AC102535 Mus muscu
29	962.5	56.0	170338	2	AC138149 Homo sapi
30	951.5	55.4	82761	9	AF391283 Homo sapi
31	950.5	55.3	948	6	AX659367 Sequence
32	936	54.5	951	10	AY317806 Sequence
33	936	54.5	236570	2	AC146610 Mus muscu
34	931.5	54.2	942	6	AX241546 Sequence
35	931.5	54.2	985	6	AX241877 Sequence
36	931.5	54.2	990	6	AX448707 Sequence
37	931.5	54.2	990	6	BD144433 Novel G-p
38	931.5	54.2	1013	6	AX702802 Sequence
39	931.5	54.2	1114	6	AX709234 Sequence
40	931.5	54.2	1364	9	AB065819 Homo sapi
41	931.5	54.2	1390	6	AX646627 Sequence
42	931.5	54.2	165729	2	AC009642 Homo sapi
43	930	54.1	235030	2	AC107331 Rattus no
44	927	54.0	954	10	AY073013 Mus muscu
45	927	54.0	954	10	AY317774 Mus muscu

ALIGNMENTS

RESULT 1

```

AC124272      185330 bp      DNA      linear      HTG 14-SEP-2002
LOCUS        Homo sapiens chromosome 11 clone RP11-607K3 map 11, WORKING DRAFT
DEFINITION   SEQUENCE, 4 unordered pieces.
AC124272      AC124272.2  GI:22857733
VERSION      HTG; HTGS_PHASE1; HTGS_DRAFT; HTGS_FULLTOP.
KEYWORDS     Homo sapiens (human)
SOURCE       Homo sapiens
ORGANISM     Homo sapiens
REFERENCE    Mammalia; Euthera; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS      Eukaryota; Euthera; Primates; Catarrhini; Hominidae; Homo.
JOURNAL      1 (bases 1 to 185330)
REFERENCE    Birren,B., Nusbaum,C. and Lander,E.
AUTHORS      Homo sapiens chromosome 11, clone RP11-607K3
JOURNAL      2 (bases 1 to 185330)
REFERENCE    Unpublished
AUTHORS      Birren,B., Linton,L., Nusbaum,C., Lander,E., Ali,A., Allen,N.,
              Anderson,S., Barna,N., Bastien,V., Bloom,T., Boguslavskiy,L.,
              Boukhgalter,B., Brown,A., Camarata,J., Campopiano,A., Chang,J.,
              Chazaro,B., Choepel,Y., Colangelo,M., Collins,S., Collymore,A.,
              Cook,A., Cooke,P., DeArellano,K., Dewar,K., Diaz,J.S., Dodge,S.,
              Faro,S., Ferreira,P., Fitzgerald,M., FitzHugh,W., Gage,D.,
              Galagan,J., Gardyna,S., Ginde,S., Gord,S., Goyette,M., Graham,L.,
              Grand-Pierre,N., Hagos,B., Horton,L., Hulme,W., Iliev,I.,
              Johnson,R., Jones,C., Kamat,A., Karatas,A., Kells,C., LaRoque,K.,
              Lamazares,R., Landers,T., Lehoczy,J., Levine,R., Lindblad-Toh,K.,
              Liu,G., MacLean,C., Macdonald,P., Major,J., Marquis,N.,
              Matthews,C., McCarthy,M., McEwan,P., McKernan,K., Meldrim,J.,
              Mihova,T., Mienna,V., Murphy,T., Naylor,J., Nguyen,C.,
              Nicol,R., Norbu,C., Norman,C.H., O'Connor,T., O'Donnell,J.,
              O'Neil,D., Oliver,J., Peterson,K., Phunkhang,P., Pierre,N.,
              Pollara,V., Raymond,C., Retta,R., Rieback,M., Riley,R., Rise,C.,
              Rogov,P., Roman,J., Rosetti,M., Roy,A., Santos,R., Schauer,S.,
              Schupback,R., Seaman,S., Severy,P., Spencer,B., Stange-Thomann,N.,
              Stojanovic,N., Strauss,N., Subramanian,A., Talamas,J., Tesfaye,S.,
              Theodore,J., Topham,K., Travers,M., Travis,N., Trigilio,J.,
              Vassiliev,H., Viel,R., Vo,A., Wilson,B., Wu,X., Wyman,D., Ye,W.J.,
              Young,G., Zainoun,J., Zembek,L., Zimmer,A. and Zody,M.
              Direct Submission
              Submitted (14-JUN-2002) Whitehead Institute/MIT Center for Genome
              Research, 320 Charles Street, Cambridge, MA 02141, USA
              3 (bases 1 to 185330)
REFERENCE    Birren,B., Nusbaum,C., Lander,E., Ali,A., Allen,N., Anderson,S.,
              Barna,N., Bastien,V., Bloom,T., Boguslavskiy,L., Boukhgalter,B.,
              Camarata,J., Chang,J., Chazaro,B., Choepel,Y., Collymore,A.,
              Cook,A., Cooke,P., DeArellano,K., Dewar,K., Diaz,J.S., Dodge,S.,
              Faro,S., Ferreira,P., Fitzgerald,M., FitzHugh,W., Gage,D., Galagan,J.,
              Gardyna,S., Gord,S., Graham,L., Grand-Pierre,N., Hagos,B.,
              Horton,L., Hulme,W., Iliev,I., Johnson,R., Jones,C., Kamat,A.,
              Karatas,A., Kells,C., Landers,T., Levine,R., Lindblad-Toh,K.,
              Liu,G., MacLean,C., Macdonald,P., Major,J., Matthews,C.,
              McCarthy,M., Meldrim,J., Mienna,V., Mienna,T., Mienna,V.,
              Murphy,T., Naylor,J., Nguyen,C., Nicol,R., Norbu,C., Norman,C.H.,
              O'Connor,T., O'Donnell,J., O'Neil,D., Oliver,J., Peterson,K.,
              Phunkhang,P., Pierre,N., Raymond,C., Retta,R., Rise,C., Rogov,P.,
              Roman,J., Roy,A., Schauer,S., Schupback,R., Seaman,S., Severy,P.,
              Smith,C., Spencer,B., Stange-Thomann,N., Stojanovic,N., Talamas,J.,
              Tesfaye,S., Theodore,J., Topham,K., Travers,M., Vassiliev,H.,
              Viel,R., Vo,A., Wilson,B., Wu,X., Wyman,D., Young,G., Zainoun,J.,
              Zembek,L., Zimmer,A. and Zody,M.
              Direct Submission
              Submitted (14-SEP-2002) Whitehead Institute/MIT Center for Genome
              Research, 320 Charles Street, Cambridge, MA 02141, USA
              On Sep 14, 2002 this sequence version replaced gi:21426295.
              All repeats were identified using RepeatMasker:
              Smit, A.F.A. & Green, P. (1996-1997)
              http://ftp.genome.washington.edu/RM/RepeatMasker.html
              ----- Genome Center
              Center: Whitehead Institute/ MIT Center for Genome Research
              Center code: WIBR
              Web site: http://www-seq.wi.mit.edu
              Contact: sequence_submissions@genome.wi.mit.edu
              ----- Project Information

```

```

Center project name: L27346
Center clone name: 607_K_3
----- Summary Statistics
Sequencing vector: Plasmid; n/a; 100% of reads
Chemistry: Dye-terminator Big Dye; 100% of reads
Assembly program: Phrap; version 0.960731
Consensus quality: 184674 bases at least Q40
Consensus quality: 184934 bases at least Q30
Consensus quality: 184980 bases at least Q20
Insert size: 180000; agarose-fp
Insert size: 185030; sum-of-contents
Quality coverage: 19.6 in Q20 bases; agarose-fp
Quality coverage: 19.1 in Q20 bases; sum-of-contents
-----
* NOTE: This is a 'working draft' sequence. It currently
* consists of 4 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.
* 1 120991: contig of 120991 bp in length
* 120992 121091: gap of 100 bp
* 121092 132589: contig of 11498 bp in length
* 132590 132689: gap of 100 bp
* 132690 153921: contig of 21232 bp in length
* 153922 154021: gap of 100 bp
* 154022 185330: contig of 31309 bp in length.
* 154022 Location/Qualifiers
* 1 .185330
* /organism="Homo sapiens"
* /mol_type="genomic DNA"
* /db_xref="taxon:9606"
* /chromosome="11"
* /map="11"
* /clone="RP11-607K3"
* /clone_lib="RPC1-11 Human Male BAC"
* 1 .120991
* /note="assembly_fragment"
* /clone_end:SP6
* /vector_side:left"
* 121092 .132589
* /note="assembly_fragment"
* 132690 .153921
* /note="assembly_fragment"
* 154022 .185330
* /note="assembly_fragment"
-----
ORIGIN
Alignment Scores:
Pred. No.: 2,68e-150 Length: 185330
Score: 1705.00 Matches: 327
Percent Similarity: 99.39% Conservative: 0
Best Local Similarity: 99.39% Mismatches: 2
Query Match: 99.24% Indels: 0
DB: 2 Gaps: 0
US-10-081-775-2 (1-329) x AC124272 (1-185330)
QY 1 MetSerSerThrLeuGlyHisAsnMetGluSerProHisHisThrAspValAspProSer 20
Db 3299 ATGTCCAGCACTTTGGCCACCAACATGGAATCTCTCATCATCATGATGTGACCTTCT 3358
QY 21 ValPhePheLeuLeuGlyIleProGlyLeuGluGlnPheHisLeuTrpLeuSerLeuPro 40
Db 3359 GTCTTCTTCTCTCTGGGCATCCAGGCTCGGAACAATTTCAATTGGGCTCTCACCCT 3418
QY 41 ValCysGlyLeuGlyThrAlaThrIleValGlyAsnIleThrIleLeuValValAla 60
Db 3419 GTGTGTGGTTAGGCACACCCACCAATTTGGGCAATATTAATCTTCTGTTGTGTGCC 3478
QY 61 ThrGluProValLeuHisLysProValTyLeuPheLeuCysMetLeuSerThrIleAsp 80

```



```

Db      3479  ACTGACCAAGCTCTTGACAGCGCTGTACCTTTCTCTGATGCTCTCACCATCGAC 3538
QY      81    LeuAlaAaSerValSerThrValProLysLeuLeuAlaIlePheTrpCysGlyAlaGly 100
Db      3539  TTGGCTGCGCTCTGTCTCCACAGTTCCCAAGCTACTGGCTATCTCTCTGGTGTGGAGCGGA 3598
QY      101   HisIleSerAlaSerAlaCysLeuAlaHisMetPhePheIleHisAlaPheCysMetMet 120
Db      3599  CATATATCTGCTCTGCTCCCTGCGCTGGCACATATGTTCTTCAATCATGCGCTTCTGCAATG 3658
QY      121   GluSerThrValLeuLeuAlaMetAlaPheAspArgTyrValAlaIleCysHisProLeu 140
Db      3659  GAGTCCACTGCTACTGCGCCATGGCTTTGATCGTAGTGGCCATCTGCCACCCACTC 3718
QY      141   ArgTyrAlaThrIleLeuThrAspThrIleLeuAlaHisIleGlyValAlaValVal 160
Db      3719  CGCTATGCGCACAACTCTCACTGACACCATCAITGGCCCATAGAGGGGTGCAGCTGTAGTG 3778
QY      161   ArgGlySerLeuLeuMetLeuProCysProPhePheIleGlyArgLeuAsnPheCysGln 180
Db      3779  CGAGGCTCCCTGCTCATGCTCCCATGCTCCCTTCTTTATTTGGCGGTGTGAACCTTCTGCCAA 3838
QY      181   SerHisValIleLeuHisThrTyrCysGluHisMetAlaValValIlysLeuAlaCysGly 200
Db      3839  AGCCATGTGATCTACACAGCTACTGTGACACATGGCTGTGGTAGCTGGCTGTGGA 3898
QY      201   AspThrArgProAsnArgValTyrGlyLeuThrAlaLeuLeuValIleGlyValAsp 220
Db      3899  GACACACAGCGCTTAACCGTGTGTATGGCTGACAGCTGCACCTGTGGTCAATGGGGTTGAC 3958
QY      221   LeuPheCysIleGlyLeuSerTyrAlaLeuIleAlaGlnAlaValLeuArgLeuSerSer 240
Db      3959  TTGTTTTCATTTGGTCTCTCCATGCTCCCTTAATGACAAAGCTGTCTTGGCTCTCATCC 4018
QY      241   HisGluAlaArgSerIysAlaLeuGlyThrCysGlySerHisValCysValIleLeuIle 260
Db      4019  CATGAAGCTCGTCTCAAGGCCCTAGGACCTGTGGTTCCTCAATGTCTGTCTATCTCATC 4078
QY      261   SerTyrThrProAlaLeuPheSerPheThrHisArgPheGlyHisHisValProVal 280
Db      4079  TCTTATACACACAGCGCTCTCTCTCTTTTACACACCGCTTTGGCCATCACGTTCCAGTC 4138
QY      281   HistIleHisIleLeuLeuAlaAsnValTyrLeuLeuLeuProProAlaLeuAsnProVal 300
Db      4139  CATATTCATCTTTTGGCCAAATGTTATCTGCTTTTGGCCACCTGCTCTTAATCCTGTG 4198
QY      301   ValTyrGlyValIysThrLysGlnIleArgLysArgValValArgValPheGlnSerGly 320
Db      4199  GTATATGAGTTAAGACCAACAGATCATGTAAGAGAGGTGTACAGGTGTTTCAAGTGGG 4258
QY      321   GlnGlyMetGlyIleLysAlaSerGlu 329
Db      4259  CAGGAATGGGCATCAAGGCATCTGAG 4285

RESULT 2
AC111179
LOCUS   AC111179
DEFINITION Homo sapiens chromosome 11 clone RP11-589G14 map 11, WORKING DRAFT
SEQUENCE, 4 ordered pieces.
ACCESSION AC111179
VERSION   AC111179.2 GI:21699469
KEYWORDS  HTG; HTGS_PHASE2; HTGS_DRAFT; HTGS_FULLTOP.
SOURCE    Homo sapiens (human)
ORGANISM  Homo sapiens
           Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
           Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 203352)
AUTHORS   Birren,B., Nusbaum,C. and Lander,E.
TITLE     Homo sapiens chromosome 11, clone RP11-589G14
JOURNAL   Unpublished
REFERENCE 2 (bases 1 to 203352)
AUTHORS   Birren,B., Linton,L., Nusbaum,C., Lander,E., Ali,A., Allen,N.,

```

```

Anderson,S., Barna,N., Bastien,V., Boguslavkiy,L., Boukhgalter,B.,
Brown,A., Camarata,J., Campopiano,A., Chang,J., Chazaro,B.,
Choepe,Y., Colangelo,M., Collins,S., Collamore,A., Cook,A.,
Cooke,P., DeArellano,K., Dewar,K., Diaz,J.S., Dodge,S., Faro,S.,
Ferreira,P., FitzHugh,W., Gage,D., Galagan,J., Gardyna,S.,
Ginde,S., Gord,S., Goyette,M., Graham,L., Grand-Pierre,N.,
Hagos,B., Horton,I., Hulme,W., Iliev,I., Johnson,R., Jones,C.,
Kamat,A., Karatas,A., Kells,C., LaRocque,K., Lamazares,R.,
Landers,T., Lehoczy,J., Levine,R., Liu,G., MacLean,C.,
Macdonald,P., Major,J., Matquis,N., Matthews,C., McCarthy,M.,
McEwan,P., McKernan,K., Meldrum,J., Meneus,L., Mihova,T.,
Mlenga,V., Murphy,T., Naylor,J., Nguyen,C., Nicol,R., Norbu,C.,
Norman,C.H., O'Connor,T., O'Donnell,P., O'Neill,D., Oliver,J.,
Peterson,K., Phunkhang,P., Pierre,N., Pollara,V., Raymond,C.,
Retta,R., Rieback,M., Riley,R., Rise,C., Rogov,P., Roman,J.,
Rosetti,M., Roy,A., Santos,R., Schauer,S., Schupback,R., Seaman,S.,
Severy,P., Spencer,B., Stange-Thomann,N., Stojanovic,N.,
Strauss,N., Subramanian,A., Talamas,J., Tesfaye,S., Theodore,J.,
Topham,K., Travers,M., Travis,N., Triglio,J., Vassiliev,H.,
Viel,R., Vo,A., Wilson,B., Wu,X., Wyman,D., Ye,W.J., Young,G.,
Zainoun,J., Zembek,L., Zimmer,A. and Zody,M.
Direct Submission
Submitted (18-FEB-2002) Whitehead Institute/MIT Center for Genome
Research, 320 Charles Street, Cambridge, MA 02141, USA
3 (bases 1 to 203352)
Birren,B., Nusbaum,C., Lander,E., Ali,A., Allen,N., Anderson,S.,
Barna,N., Bastien,V., Bloom,T., Boguslavkiy,L., Boukhgalter,B.,
Camarata,J., Chang,J., Chazaro,B., Choepe,Y., Collamore,A.,
Cooke,A., Cooke,P., DeArellano,K., Dewar,K., Diaz,J.S., Dodge,S.,
Faro,S., Ferreira,P., FitzGerald,M., Gage,D., Galagan,J.,
Gardyna,S., Gord,S., Graham,L., Grand-Pierre,N., Hagos,B.,
Horton,L., Hulme,W., Iliev,I., Johnson,R., Jones,C., Kamat,A.,
Karatas,A., Kells,C., Landers,T., Levine,R., Lindblad-Toh,K.,
Liu,G., MacLean,C., Macdonald,P., Major,J., Matthews,C.,
McCarthy,M., Meldrum,J., Meneus,L., Mihova,T., Mlenga,V.,
Murphy,T., Naylor,J., Nguyen,C., Nicol,R., Norbu,C., Norman,C.H.,
O'Connor,T., O'Donnell,P., O'Neill,D., Oliver,J., Peterson,K.,
Phunkhang,P., Pierre,N., Raymond,C., Retta,R., Rise,C., Rogov,P.,
Roman,J., Roy,A., Schauer,S., Schupback,R., Seaman,S., Severy,P.,
Smith,C., Spencer,B., Stange-Thomann,N., Stojanovic,N., Talamas,J.,
Tesfaye,S., Theodore,J., Topham,K., Travers,M., Vassiliev,H.,
Viel,R., Vo,A., Wilson,B., Wu,X., Wyman,D., Young,G., Zainoun,J.,
Zembek,L., Zimmer,A. and Zody,M.
Direct Submission
Submitted (06-AUG-2002) Whitehead Institute/MIT Center for Genome
Research, 320 Charles Street, Cambridge, MA 02141, USA
On Jul 5, 2002 this sequence version replaced gi:18695944.
All repeats were identified using RepeatMasker:
Smit, A.F.A. & Green, P. (1996-1997)
http://ftp.genome.washington.edu/RM/RepeatMasker.html
-----
Center: Whitehead Institute/ MIT Center for Genome Research
Center code: WIBR
Web site: http://www-seq.wi.mit.edu
Contact: sequence_submissions@genome.wi.mit.edu
-----
Project Information
Center project name: L25440
Center clone name: 589 G14
-----
Summary Statistics
Sequencing vector: Plasmid; n/a; 100% of reads
Chemistry: Dye-terminator Big Dye; 100% of reads
Assembly program: Phrap; version 0.960731
Consensus quality: 201894 bases at least Q40
Consensus quality: 202347 bases at least Q30
Consensus quality: 202672 bases at least Q20
Insert size: 200000; agarose-fp
Insert size: 203052; sum-of-ctngs
Quality coverage: 14.6 in Q20 bases; agarose-fp
Quality coverage: 14.4 in Q20 bases; sum-of-ctngs
-----
* NOTE: This is a 'working draft' sequence. It currently
* consists of 4 contigs. Gaps between the contigs
* are represented as runs of N. The order of the pieces

```

TITLE
JOURNAL
REFERENCE
AUTHORS

TITLE
JOURNAL
COMMENT

* is believed to be correct as given, however the sizes
 * of the gaps between them are based on estimates that have
 * provided by the submitter.
 * This sequence will be replaced
 * by the finished sequence as soon as it is available and
 * the accession number will be preserved.

* 1 70426: contig of 70426 bp in length
 * 70427 70526: gap of 100 bp
 * 70527 78730: contig of 8204 bp in length
 * 78731 78830: gap of 100 bp
 * 78831 137440: contig of 58610 bp in length
 * 137441 137540: gap of 100 bp
 * 137541 203352: contig of 65812 bp in length.

FEATURES

source

1..203352
 /organism="Homo sapiens"
 /mol_type="genomic DNA"
 /db_xref="taxon:9606"
 /chromosome="11"
 /map="11"
 /clone="RP11-589G14"
 /clone_lib="RPC1-11 Human Male BAC"

misc_feature

1..70426
 /note="assembly_fragment
 clone_end:SP6
 vector_side:left"

misc_feature

70527..78730
 /note="assembly_fragment"
 78831..137440

misc_feature

/note="assembly_fragment"
 137541..203352

misc_feature

/note="assembly_fragment
 clone_end:P7
 vector_side:right"

ORIGIN

Alignment Scores:
 Pred. No.: 2,99e-150 Length: 203352
 Score: 1705.00 Matches: 327
 Percent Similarity: 99.39% Conservative: 0
 Best Local Similarity: 99.39% Mismatches: 2
 Query Match: 99.24% Indels: 0
 DB: 2 Gaps: 0

US-10-081-775-2 (1-329) x AC111179 (1-203352)

QY 1 MetSerThrLeuGlyHisAsnMetGluSerProHisHisThrAspValAspProSer 20
 DB 169277 ATGTCAGCACTCTTGGCCACAAACATGGAATCTCCTCATCACATGATGTTGACCTTCT 169336

QY 21 ValPheLeuLeuGlyIleProGlyLeuGluGlnPheHisLeuTrpLeuSerLeuPro 40
 DB 169337 GTCCTTCTCTCTGGGATCCAGGCTGGAAACAAATTCATTGTGGCTCTCACTCCCT 169396

QY 41 ValCysGlyLeuGlyThrAlaThrIleValGlyAsnIleThrIleLeuValValAla 60
 DB 169397 GGTGTGGCTTAGGCACACCAATTTGGGCAATATAACTATTCTGTGTGTGCC 169456

QY 61 ThrGluProValLeuHisLysProValTyrLeuPheLeuCysMetLeuSerThrIleAsp 80
 DB 169457 ACTGAACCACTTGGCACAAGCTGTGTACCTTTTCTGTGATGCTCTCAACCACTGAC 169516

QY 81 LeuAlaAlaSerValSerThrValProLysLeuLeuAlaIlePheTrpCysGlyAlaGly 100
 DB 169517 TTGGCTGCTCTGTCTCCACAGTTCACAGCTACTGGCTATCTTCTGGTGGAGCCGGA 169576

QY 101 HisIleSerAlaSerAlaCysLeuAlaHisMetPhePheIleHisAlaPheCysMetMet 120
 DB 169577 CATATATCTGCTCTGCTGCTGCTGGCACAATATGTTCTTCATTCATGCTTCTGCATGATG 169636

QY 121 GluSerThrValLeuLeuAlaMetAlaPheAspArgTyrValAlaIleCysHisProLeu 140
 DB 169637 GAGTCACCTGTGCTACTGTGCCCATGGCTTTGTATGCTAGCTGGCCATCTGCCCACTC 169696

QY 141 ArgTyrAlaThrIleLeuThrAspThrIleAlaHisIleGlyValAlaValVal 160
 DB 169697 CGCTATGCCACAATCCTCCTACCTGACACCAATATTGCCACATAGGGGTGGAGCTGTAGTG 169756

QY 161 ArgGlySerLeuLeuMetLeuProCysProPhePheIleGlyArgLeuAsnPheCysGln 180
 DB 169757 CGAGGCTCCCTGCTCATGCTCCCATGTCCTTTTATTGGCGGTTTGAACCTTCTGCCAA 169816

QY 181 SerHisValIleLeuHisThrTyrCysGluHisMetAlaValValLysLeuAlaCysGly 200
 DB 169817 AGCCATGTGATCTTACACACGCTACTGTGAGCACATGGCTGTGGTGAAGCTGCCCTGTGGA 169876

QY 201 AspThrArgProAsnArgValTyrGlyLeuThrAlaAlaLeuLeuValIleGlyValAsp 220
 DB 169877 GACACCAAGCTTACCCGTTGATGGGCTGACAGCTGACCTGTGGTTCATTTGGGCTTGAC 169936

QY 221 LeuPheCysIleGlyLeuSerTyrAlaLeuLeuAlaGlnAlaValLeuArgLeuSerSer 240
 DB 169937 TTGTTTTGCATTGGTCTCTCCTATGCTTAAATGGCAAGCTGTCTTCTGCCCTCTCATCC 169996

QY 241 HisGluAlaArgSerLysAlaLeuGlyThrCysGlySerHisValCysValIleLeuIle 260
 DB 169997 CATGAAGCTCGTCCAGGCTTAGGACCTGTGGTCTCCCATGTGTGTCTCATCTCATC 170056

QY 261 SerTyrThrProAlaLeuPheSerPheThrHisArgPheGlyHisHisValProVal 280
 DB 170057 TCTTATACACAGCCCTCTCTCTTTTACACACCGCTTTGGCCATCAGTTCAGTTC 170116

QY 281 HisIleHisIleLeuLeuAlaAsnValTyrLeuLeuLeuProProAlaLeuAsnProVal 300
 DB 170117 CATATTCACTTCTTTTGGCCAAATGTTTATCTCTCTTTGGCCACCTGCTTAACTCTGTG 170176

QY 301 ValTyrGlyValLysThrLysGlnIleArgLysArgValValArgValPheGlnSerGly 320
 DB 170177 GTATATGAGTTAAGACCAACACATCAGTAAAGAGGTGTGAGGTGTTTCAAGAGTGGG 170236

QY 321 GlnGlyMetGlyIleLysAlaSerGlu 329

DB 170237 CAGGGAATGGGCATCAGGCATCTGAG 170263

RESULT 3
 BD144452
 LOCUS BD144452 990 bp DNA linear PAT 17-JAN-2003
 DEFINITION Novel G-protein coupled receptors.
 ACCESSION BD144452
 VERSION BD144452.1 GI:27850210
 KEYWORDS JP 2002112793-A/177.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini, Hominidae; Homo.
 REFERENCE 1 (bases 1 to 990)
 AUTHORS Haga,T., Takeda,S. and Miyake,N.
 TITLE Novel G-protein coupled receptors
 JOURNAL Patent: JP 2002112793-A 177 16-APR-2002;
 JAPAN SCIENCE AND TECHNOLOGY CORP
 COMMENT OS Homo sapiens (human)
 FN JP 2002112793-A/177
 PD 16-APR-2002
 PF 03-FEB-2001 JP 2001034434
 PI TATSUYA HAGA, SHIGEKI TAKEDA, NARIKI MIYAKE
 PC C12N15/09,A01K67/027,A61K38/00,A61K39/395,A61K39/395,A61K45/00, PC
 A61K48/00,
 PC A61P43/00,C07K14/705,C07K16/28,C07K19/00,C12N1/15,C12N1/19, PC
 C12N1/21,
 PC C12N5/10,C12Q1/02,C12Q1/68,G01N33/15,G01N33/50,G01N33/53, PC
 G01N33/566//
 PC C12P21/08,C12N15/00,A61K37/02,C12N5/00
 CC Novel G-protein coupled receptors
 FT Key Location/Qualifiers
 FT CDS (1)..(990).

```

FEATURES
    source
        Location/Qualifiers
            1..990
                /organism="Homo sapiens"
                /mol_type="genomic DNA"
                /db_xref="taxon:9606"

ORIGIN
Alignment Scores:
Pred. No.: 1..31e-151 Length: 990
Score: 1691.00 Matches: 325
Percent Similarity: 99.09% Conservative: 1
Best Local Similarity: 98.78% Mismatches: 3
Query Match: 98.43% Indels: 0
DB: 6 Gaps: 0

US-10-081-775-2 (1-329) x BD144452 (1-990)
QY 1 MetSerSerThrLeuGlyHisAsnMetGluSerProHisHisThrAspValAspProSer 20
DB 1 ATGTCAGACACTCTTGGCCACACATGGAATCTCCTAATCACACTGATGTTGACCCCTTCT 60
QY 21 ValPhePheLeuGlyIleProGlyLeuGluGlnPheHisLeuTrpLeuSerLeuPro 40
DB 61 GTCTTCTCTCTCTGCGCATCCAGGCTCGGACCAATTTCAATTTGTTGGCTCTCACTCCCT 120
QY 41 ValCysGlyLeuGlyThrAlaThrIleValGlyAsnIleThrIleLeuValValAla 60
DB 121 GGTGTGGCTTAGGCACACCAATGTTGGCAATATTAACATATTTCTGTTGTTGTGTC 180
QY 61 ThrGluProValLeuHisHisProValTyrLeuPheLeuPheCysMetLeuSerThrIleAsp 80
DB 181 ACTGAACAGCTCTTGCAAGCTGTGTACCTTTTCTGTGATGCTCTCAACCATCGAC 240
QY 81 LeuAlaAlaSerValSerThrValProLysLeuLeuAlaIlePheTrpCysGlyAlaGly 100
DB 241 TTGGCTGCTCTGTCTCCACAGTCTCCCAAGCTACTGGCTATCTTCTGGTGTGGAGCCGA 300
QY 101 HisIleSerAlaSerAlaCysLeuAlaHisMetPhePheIleHisAlaPheCysMetMet 120
DB 301 CATATATCTGCTCTGCTGCTGCGCACAGATGTTCTTCAATTCATGCTCTTGCATGATG 360
QY 121 GluSerThrValLeuLeuAlaMetAlaPheAspArgTyrValAlaIleCysHisProLeu 140
DB 361 GAGTCCACTGTCTACTTGCCATGCTCTTGTATGCTAGTGGCCATCTGCCACCCACTC 420
QY 141 ArgTyrAlaThrIleLeuThrAspThrIleIleAlaHisIleGlyValAlaAlaValVal 160
DB 421 CGCTATGCCAATCTCTCACTGACACCATCATTTGCCCAATAGGGGCTGGCAGCTTAGTG 480
QY 161 ArgGlySerLeuLeuMetLeuProCysProPhePheIleGlyArgLeuAsnPheCysGln 180
DB 481 CGAGGCTCCCTGCTCATGCTCCCATGCTCCCTTCTTATTTGGCGTTGAACTTCTGCCAA 540
QY 181 SerHisValIleLeuHisThrTyrCysGluHisMetAlaValValLysLeuAlaCysGly 200
DB 541 AGCCATGTGATCCTATACACAGTACTGTGAGCACATGGCTGTGGTGAAGCTGCGCTGTGA 600
QY 201 AspThrArgProAsnArgValTyrGlyLeuThrAlaAlaLeuLeuValIleGlyValAsp 220
DB 601 GACACAGGCTTAACCGGTGTATGGGTGACAGCTGCATGCTGTTGGTCAATTTGGGTTTAC 660
QY 221 LeuPheCysIleGlyLeuSerTyrAlaLeuIleAlaGlnAlaValLeuArgLeuSerSer 240
DB 661 TTGTTTTCATTTGGTCTCTCTATGCTCTTCTAGTGCACAGCTGCTCTTGGCTCTCATCC 720
QY 241 HisGluAlaArgSerLysAlaLeuGlyThrCysGlySerHisValCysValIleLeuIle 260
DB 721 CATGAAGCTCGGCTCAAGGCTTAGGACCTGTGGTTCCCATGCTCTGTTGATCCTCATC 780
QY 261 SerTyrThrProAlaLeuPheSerPhePheThrHisArgPheGlyHisHisValProVal 280
DB 781 TCTTATACACGAGCCCTTCTCTCTTTTACACACCGCTTTGGCCATCAGCTTCCAGTC 840

```

```

QY 281 HisIleHisIleLeuLeuAlaAsnValTyrLeuLeuLeuProProAlaLeuAsnProVal 300
DB 841 CATATTACATCTCTTTGGCCAAATGTTATCTGCTTTTGCCACCTGCTCTTAATCTCTGTG 900
QY 301 ValTyrGlyValLysThrLysGlnIleArgLysArgValValArgValPheGlnSerGly 320
DB 901 GTATATGGAGTTAAGACCAACAGATCCGTAAGAAGATTGTCAAGGTGTTTCAAAGTGGG 960
QY 321 GlnGlyMetGlyIleLysAlaSerGlu 329
DB 961 CAGGGAATGGCATCAAGGCATCTGAG 987
RESULT 4
AX551417
LOCUS AX551417
DEFINITION Sequence 35 from Patent WO0250276.
ACCESSION AX551417
VERSION AX551417.1 GI:25814217
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1
AUTHORS Li,L., Padigaru,M., Ballinger,R.A., Kekuda,R., Colman,S.D.,
Sciore,P., Smithson,G., Peyman,J.A., Macdougall,J.R., Stone,D.,
Vernet,C.A., Shenoy,S., Gunther,E., Millet,I., Tchernev,V.T.,
Anderson,D., Gusev,V., Malyankar,U.M., Zhong,H., Ellerman,K.E. and
Wolenc,A.
TITLE Novel proteins and nucleic acids encoding same
JOURNAL Patent: WO 0250276-A 35 27-JUN-2002;
Curagen Corporation (US)
FEATURES
    source
        1..998
            /organism="Homo sapiens"
            /mol_type="unassigned DNA"
            /db_xref="taxon:9606"
        2..991
            /note="unnamed protein product"
            /codon_start=1
            /protein_id="CAD57920.1"
            /db_xref="GI:25814218"
            /db_xref="REMBL:CAD57920"
            /translation="MSSSLGHNSPNHTDVPVFLFLGIPGLEQFHLWLSPLVCGL
            GTATVGNITLIVVATEPVLHKPVLKMLSTIDLAASVSTPKLAIIPWCGAGHI
            SASCLAQMFIFHAFCMMESTVLLAMAFDRYVAICHPLRYATILDTITIAHIGAAVH
            RGLMLPCPELIGRLNFCQSHVILHTYCEHMAVVKLACGTRPNRVGLTAAVLVIG
            VDLFCIGLSYALSAQAVLRLLSSHEARSKALCTCGSHVCVILISYTPALFSPETHRFGH
            HVPVHIHLANVILLPPLNPVYGVTKQIRKRVVRVTFQSGGGMGIKASE"
ORIGIN
Alignment Scores:
Pred. No.: 1..32e-151 Length: 998
Score: 1691.00 Matches: 325
Percent Similarity: 99.09% Conservative: 1
Best Local Similarity: 98.78% Mismatches: 3
Query Match: 98.43% Indels: 0
DB: 6 Gaps: 0

US-10-081-775-2 (1-329) x AX551417 (1-998)
QY 1 MetSerSerThrLeuGlyHisAsnMetGluSerProHisHisThrAspValAspProSer 20
DB 2 ATGTCCAGCACTCTTGGCCACACATGGAATCTCCTAATCACACTGATGTTGACCCCTTCT 61
QY 21 ValPhePheLeuGlyIleProGlyLeuGluGlnPheHisLeuTrpLeuSerLeuPro 40
DB 62 GTCTTCTCTCTCTGCGCATCCAGGCTCGGACCAATTTCAATTTGTTGGCTCTCACTCCCT 121
QY 41 ValCysGlyLeuGlyThrAlaThrIleValGlyAsnIleThrIleLeuValValAla 60
DB 122 GTGTGTGGCTTAGGCACACACCAATTTGGCCAAATATTAACATATTTCTGTTGTTGTGTC 181

```

```
QY 61 ThrGluProValLeuHisIysProValTyrLeuPheLeuCysMetLeuSerThrIleAsp 80
DB 182 ACTGAACAGCTGTGCACAAAGCTGTACCTTTTCTGTGCATGCTCTCAACCATCGAC 241

QY 81 LeuAlaAlaSerValSerThrValProLysLeuLeuAlaIlePheTyrCysGlyAlaGly 100
DB 242 TTGGCTGCTGTCTCCACAGTTCACAGCTACTGGCTAATCTTCTGTGTGGAGCCGA 301

QY 101 HisIleSerAlaSerAlaCysLeuAlaHisMetPhePheIleHisAlaPheCysMetMet 120
DB 302 CATATATCTGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 361

QY 121 GluSerThrValLeuAlaMetAlaPheAspArgTyrValAlaIleCysHisProLeu 140
DB 362 GAGTCCACTGTGTACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 421

QY 141 ArgTyrAlaThrIleLeuThrAspThrIleAlaHisIleGlyValAlaAlaValVal 160
DB 422 CGCTATGCCAATCTCTCACTGACCATCATTTGCCACATAGGGGTGGCAGCTGTAGTG 481

QY 161 ArgGlySerLeuLeuMetLeuProCysProPhePheIleGlyArgLeuAsnPheCysGln 180
DB 482 CGAGGCTCCCTGCTCATGCTCCATGCTCCCTTCTTATTTGGGCGTTTGAACCTTCCCAA 541

QY 181 SerHisValIleLeuHisThrTyrCysGluHisMetAlaValValIleCysGly 200
DB 542 AGCCATGTGATCTTACACAGCTACTGTGACCATGCTGCTGCTGCTGCTGCTGCTGCTG 601

QY 201 AspThrArgProAsnArgValTyrGlyLeuThrAlaAlaLeuValIleGlyValAsp 220
DB 602 GACACAGGCTTAACCGTGTGTATGGCTGACAGCTGCTGCTGCTGCTGCTGCTGCTGCTG 661

QY 221 LeuPheCysIleGlyLeuSerTyrAlaLeuIleAlaGlnAlaValLeuArgLeuSerSer 240
DB 662 TTGTTTTGCAATGGTCTCTCTATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 721

QY 241 HisGluAlaArgSerIysAlaLeuGlyThrCysGlySerHisValCysValIleLeuIle 260
DB 722 CATGAAGCTCGGTCCAGGCCCTTAGGACCTGTGTGCTTCCATGTCGTGCTGCTGCTGCTG 781

QY 261 SerTyrThrProAlaLeuPheSerPhePheThrHisArgPheGlyHisHisValProVal 280
DB 782 TCTATACACAGGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 841

QY 281 HisIleHisIleLeuAlaSerValTyrLeuLeuLeuProProAlaLeuAsnProVal 300
DB 842 CATATTCATCTCTTTTGGCCAAATGTTATCTGCTTTTGCCACCTGCTCTTAATCTGTG 901

QY 301 ValTyrGlyValLysThrLysGlnIleArgLysArgValValArgValPheGlnSerGly 320
DB 902 GTATATGGAGTTAAGACCAACAGATCCGTAAAGAGTTGTAGGGTGTTCAGAGTGGG 961

QY 321 GlnGlyMetGlyIleLysAlaSerGlu 329
DB 962 CAGGGAATGGGCATCAAGGCACTCGAG 988
```

RESULT 5
AX551419
LOCUS
DEFINITION
Sequence 37 from Patent WO0250276.
ACCESSION
AX551419
VERSION
AX551419.1 GI:25814219
KEYWORDS
Homo sapiens (human)
SOURCE
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE
1
Li, L., Padigar, M., Ballinger, R.A., Kekuda, R., Colman, S.D.,
Sciore, P., Smithson, G., Peyton, J.A., MacDougall, J.R., Stone, D.,
Vernet, C.A., Shenoy, S., Gunther, E., Millet, I., Tchernev, V.T.,
Anderson, D., Gusev, V., Malyankar, U.M., Zhong, H., Ellerman, K.E. and
Wolenc, A.

TITLE Novel proteins and nucleic acids encoding same
JOURNAL Patent: WO 0250276-A 37 27-JUN-2002;
Curagen Corporation (US)
FEATURES Location/Qualifiers
source
1. .998
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"
2. .991
/note="unnamed protein product"
/codon_start=1
/protein_id="CAD57921.1"
/db_xref="GI:25814220"
/translation="MSSLTGHNMBSPNHTDVPVFFLLGIPGLBQPHLMLSLPYCGGL
GTATVGNITILVVVATEPVLUHPVYFLCMLSTIDLAASVTPVKLLAIWCGAGHI
SASCAQMFHFAFCMMESVLLAMAFDRAVVAICHPRYATILDTTIIAHIGAAVIV
RGLMLPFCPLIQRLNQSHVILHTYCEHMAVVKACGDRNVRVGLAALVIG
VDLCFGLSVALSAQVLRSLSHSARSALGTCGSHVGVILISYTPALFSFTFRFGH
HVPVHIHLLANVYLLPPALNPVYVGVTKQIRKRVVRVPSGQGMGKASE"

CDS
/note="unnamed protein product"
/codon_start=1
/protein_id="CAD57921.1"
/db_xref="GI:25814220"
/translation="MSSLTGHNMBSPNHTDVPVFFLLGIPGLBQPHLMLSLPYCGGL
GTATVGNITILVVVATEPVLUHPVYFLCMLSTIDLAASVTPVKLLAIWCGAGHI
SASCAQMFHFAFCMMESVLLAMAFDRAVVAICHPRYATILDTTIIAHIGAAVIV
RGLMLPFCPLIQRLNQSHVILHTYCEHMAVVKACGDRNVRVGLAALVIG
VDLCFGLSVALSAQVLRSLSHSARSALGTCGSHVGVILISYTPALFSFTFRFGH
HVPVHIHLLANVYLLPPALNPVYVGVTKQIRKRVVRVPSGQGMGKASE"

ORIGIN

Alignment Scores:
Pred. No.: 1.32e-151 Length: 998
Score: 1691.00 Matches: 325
Percent Similarity: 99.0% Conservative: 1
Best Local Similarity: 98.7% Mismatches: 3
Query Match: 98.4% Indels: 0
DB: 6 Gaps: 0
US-10-081-775-2 (1-329) x AX551419 (1-998)

QY 1 MetSerSerThrLeuGlyHisAsnMetGluSerProHisHisThrAspValAspProSer 20
DB 2 ATGTCCAGACCTCTTGGCCACCAACATGGAATCTCTAATACACATGATGTGACCTTCT 61
QY 21 ValPhePheLeuLeuGlyLeuProGlyLeuGluGlnPheHisLeuTrpLeuSerLeuPro 40
DB 62 GTCTTCTTCTCTGCGCATCCAGTCTGGAAACAATTCATTTTGGTCTCTCACTCCCT 121
QY 41 ValCysGlyLeuGlyThrAlaThrIleValGlyAsnIleThrIleLeuValValAla 60
DB 122 GTGTGTGGCTTAGGCACAGCCACAATTTGGGCAATATAAATTAATTTCTGTTGTGTC 181
QY 61 ThrGluProValLeuHisIysProValTyrLeuPheLeuCysMetLeuSerThrIleAsp 80
DB 182 ACTGAACAGCTGTGCACAAAGCTGTACCTTTTCTGTGCATGCTCTCAACCATCGAC 241
QY 81 LeuAlaAlaSerValSerThrValProLysLeuLeuAlaIlePheTyrCysGlyAlaGly 100
DB 242 TTGGCTGCTGTGTCTCCACAGTTCACAGCTTCCCAAGCTACTGGCTATCTTCTGTGTG 301
QY 101 HisIleSerAlaSerAlaCysLeuAlaHisMetPhePheIleHisAlaPheCysMetMet 120
DB 302 CATATATCTGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 361
QY 121 GluSerThrValLeuAlaMetAlaPheAspArgTyrValAlaIleCysHisProLeu 140
DB 362 GAGTCCACTGTGTACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 421
QY 141 ArgTyrAlaThrIleLeuThrAspThrIleAlaHisIleGlyValAlaAlaValVal 160
DB 422 CGCTATGCCAATCTCTCACTGACCATCATTTGCCACATAGGGGTGGCAGCTGTAGTG 481
QY 161 ArgGlySerLeuLeuMetLeuProCysProPhePheIleGlyArgLeuAsnPheCysGln 180
DB 482 CGAGGCTCCCTGCTCATGCTCCATGCTCCCTTCTTATTTGGGCGTTTGAACCTTCCCAA 541
QY 181 SerHisValIleLeuHisThrTyrCysGluHisMetAlaValValIleCysGly 200
DB 542 AGCCATGTGATCTTACACAGCTACTGTGACCATGCTGCTGCTGCTGCTGCTGCTGCTG 601
QY 201 AspThrArgProAsnArgValTyrGlyLeuThrAlaAlaLeuValIleGlyValAsp 220

```

Db      602  GACACAGGCGCTAAACCGTGTGTATGGCTGACAGCTGCATGTGTGCTCATTTGGGGTTGAC 661
QY      221  LeuPheCysIleGlyLeuSerTyrAlaLeuIleAlaGlnAlaValLeuArgLeuSerSer 240
Db      662  TTGTTTTCATTTGGCTCTCTCTCTATGCCCTAAAGTGCACAAGCTGCTCTGCGCTCTCATCC 721
QY      241  HisGluAlaArgSerLysAlaLeuGlyThrCysGlySerHisValCysValIleLeuIle 260
Db      722  CATGAAGCTCGGTCCAGGCGCTTAGGGACCTGTGGTTCCTCATGCTGTGTCTCATCTCATC 781
QY      261  SerTyrThrProAlaLeuPheSerPheThrHisArgPheGlyHisHisValProVal 280
Db      782  TCTTATACACAGCGCTCTCTCTCTCTTTTACACACCGCTTTGGCCATCACGTTCCAGTC 841
QY      281  HisIleHisIleLeuLeuAlaSerValTyrLeuLeuLeuProProAlaLeuAsnProVal 300
Db      842  CATATTTCATCTTTTGGCCAAATGTTTATCTGCTTTTGGCCACCTGCTCTTAATCTCTGTG 901
QY      301  ValTyrGlyValLysThrLysGlnIleArgLysArgValValAlaValArgValPheGlnSerGly 320
Db      902  GTATATGGAGTTAAGACCAACAGATCCCTAAAGAGTTGTGAGGTTGTTTCAAAGTGGG 961
QY      321  GlnGlyMetGlyIleLysAlaSerGlu 329
Db      962  CAGGGAATGGGCATCAAGCACTCTGAG 988

RESULT 6
AX646605
LOCUS      AX646605      1390 bp      DNA      linear      PAT 04-MAR-2003
DEFINITION      Sequence 797 from Patent EP1270724.
ACCESSION      AX646605
VERSION      AX646605.1  GI:28798988
KEYWORDS
SOURCE      Homo sapiens (human)
ORGANISM      Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE      1
AUTHORS      Suwa, M., Asai, K., Akiyama, Y. and Aburatani, H.
TITLE      Guanosine triphosphate-binding protein coupled receptors
JOURNAL      Patent: EP 1270724-A 797 02-JAN-2003;
National Institute of Advanced Industrial Science and Technology
(JP) ; Center for Advanced Science and Technology Incubation, Ltd.
(JP)

FEATURES             Location/Qualifiers
     source           1..1390
                     /organism="Homo sapiens"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:9606"
     CDS               201..1190
                     /note="unnamed protein product"
                     /codon_start=1
                     /protein_id="CAD69429.1"
                     /db_xref="GI:28798989"
                     /db_xref="REMBL:CAD69429"
                     /translation="MSSTLGHMSPHNDVDPVFFLLGIPGLEQLHMLSLPVCGL
GTATIVGNILVVFVPEVLPKPVYFLFLCMLSTDLAASVSTPKLLAIPWCGAGHI
SASCLAQMFIFHAFWMEVTLAMAFDYVAICHELRYATILDTIIAHIGVAAYV
RGSLLMLPCPLIGRLNFCQSHVLIHTYCEHMAVVKLACGDRPNRVGLTAALLVIG
VDLFCIGLSVALSAQAVRLSSHEARSKALGTCGSHVCVILISVTPALFSPFTRFGH
HVPVHIILLANVYLLLPALNPVYGVKTKQIRKVRVVFQSGQGMGIKASE"

ORIGIN
Alignment Scores:
Pred. No.:      1.94e-151      Length:      1390
Score:          1691.00      Matches:    325
Percent Similarity: 99.09%      Conservative: 1
Best Local Similarity: 98.78%      Mismatches: 3
Query Match:    98.43%      Indels:      0
DB:             6      Gaps:      0

US-10-081-775-2 (1-329) x AX646605 (1-1390)

```

```

QY      1  MetSerSerThrIleuGlyHisAsnMetGluSerProHisHisThrAspValAspProSer 20
Db      201  ATGTCAGCACCTTTGGCCACAACATGGAATCTCTTAATCACACATGATGTTGACCTTCT 260
QY      21  ValPhePheLeuLeuGlyIleProGlyLeuGluGlnPheHisLeuTrpLeuSerLeuPro 40
Db      261  GTCTTCTCTCTCTGCGCATCCAGGCTCGAACAATTTCAATTGTGGCTCTCACTCCCT 320
QY      41  ValCysGlyLeuGlyThrAlaThrIleValGlyAsnIleThrIleLeuValValAla 60
Db      321  GTGTGTGGCTTAGGCACAGCCCAATTTGGGCAATATTAACCTATTCTGGTGTGTGTGCC 380
QY      61  ThrGluProValLeuHisLysProValTyrLeuPheLeuCysMetLeuSerThrIleAsp 80
Db      381  ACTGAACCAAGCTCTGCACAAGCTGTGTACCTTTTCTGTGCATGCTCTCAACCATCGAC 440
QY      81  LeuAlaAlaSerValSerThrValProLysLeuLeuAlaIlePheTrpCysGlyAlaGly 100
Db      441  TTGGCTGCGCTCTGTCTCCACAGTTCCCAAGCTACTGGCTATCTTCTGTGTGTGGAGCGGA 500
QY      101  HisIleSerAlaSerAlaCysLeuAlaHisMetPhePheIleHisAlaPheCysMetMet 120
Db      501  CATATATCTGCTCTGCTGCTGCTGCGCACAGATGTTCTTCAITCATGCTTTCGANGATG 560
QY      121  GluSerThrValLeuLeuAlaMetAlaPheAspArgTyrValAlaIleCysHisProLeu 140
Db      561  GAGTCCACTGTCTACTGGCCATGGCTTTGATCGCTAGCTGGCCATCTGCCACCCACTC 620
QY      141  ArgTyrAlaThrIleLeuThrAspThrIleIleAlaHisIleGlyValAlaAlaValVal 160
Db      621  CGCTATGCGCAATCTCTCACTGACACCATCATATGCCACATAGGGGTGGCAGCTGTAGTG 680
QY      161  ArgGlySerLeuLeuMetLeuProCysProPhePheIleGlyArgLeuAsnPheCysGln 180
Db      681  CGAGCTCCCTGCTCATGCTCCCATGCTCTCTTATGGCGCTTGAACCTTCTGCCAAC 740
QY      181  SerHisValIleLeuHisThrTyrCysGluHisMetAlaValValLysLeuAlaCysGly 200
Db      741  AGCCATGTGATCCTACACAGCTACTGTGAGCACATGCTGTGGTGAAGCTGGCTGTGGA 800
QY      201  AspThrArgProAsnArgValTyrGlyLeuThrAlaAlaLeuLeuValIleGlyValAsp 220
Db      801  GACACAGGCGCTAAACCGTGTGTATGGGTGACAGTGCACCTGTTGGTCAITGGGTTGAC 860
QY      221  LeuPheCysIleGlyLeuSerTyrAlaLeuIleAlaGlnAlaValLeuArgLeuSerSer 240
Db      861  TTGTTTTCATTTGGTCTCTCTATGCTTAACTAGTGCACAGCTGCTTTCGGCTCTCATCC 920
QY      241  HisGluAlaArgSerLysAlaLeuGlyThrCysGlySerHisValCysValIleLeuIle 260
Db      921  CATGAAGCTCGGTCCAGGCGCTTAGGACCTGTGGTTCCTCCATGCTGTGTCATCTCATC 980
QY      261  SerTyrThrProAlaLeuPheSerPhePheThrHisArgPheGlyHisHisValProVal 280
Db      981  TCTTATACACAGCGCTCTCTCTCTCTTTTACACACCGCTTTGGCCATCACGTTCCAGTC 1040
QY      281  HisIleHisIleLeuLeuAlaAsnValTyrLeuLeuLeuProProAlaLeuAsnProVal 300
Db      1041  CATATTTCATCTTTTGGCCAAATGTTTATCTGCTTTTGGCCACCTGCTCTTAATCTCTGTG 1100
QY      301  ValTyrGlyValLysThrLysGlnIleArgLysArgValValAlaValPheGlnSerGly 320
Db      1101  GTATATGGAGTTAAGACCAACAGATCCGTAAGAGAGTTGTGAGGAGTTTCAAAGTGGG 1160
QY      321  GlnGlyMetGlyIleLysAlaSerGlu 329
Db      1161  CAGGGAATGGGCATCAAGCATCTGAG 1187

RESULT 7
AB065534
LOCUS      AB065534      1390 bp      DNA      linear      PRI 23-JUL-2002
DEFINITION      Homo sapiens gene for seven transmembrane helix receptor, complete

```

cds, isolate:CBRC7TM_97. AB065534 AB065534.1 GI:21928364 SOURCE Homo sapiens (human) ORGANISM Homo sapiens Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo. 1 Suwa, M., Sato, T., Okouchi, I., Arita, M., Futami, K., Matsumoto, S., Tsutsumi, S., Aburatani, H., Asai, K. and Akiyama, Y. Genome-wide discovery and analysis of human seven transmembrane helix receptor genes Unpublished 2 (bases 1 to 1390) Suwa, M. Direct Submission Submitted (11-JUL-2001) Makiko Suwa, Computational Biology Research Center (CBRC), National Institute of Advanced Industrial Science and Technology (AIST); 2-41-6 Aomi Koto-ku, Tokyo 135-0064, Japan (E-mail: m-suwa@aist.go.jp; URL: http://www.cbrc.jp/; Tel: 81-3-3599-8080, Fax: 81-3-3599-8081) This sequence is a seven transmembrane helix receptor candidate predicted from the whole human genome sequences using our automated system that contains programs of gene finding (GeneDecoder), sequence search, motif-domain assignment and transmembrane helix prediction. And the sequence is submitted by the collaborative project between [Computational Biology Research Center (CBRC), National Institute of Advanced Industrial Science and Technology (AIST)] and [Genome Science Division, Research Center for Advanced Science and Technology (RCAT), University of Tokyo]. Location/Qualifiers 1..1390 /organism="Homo sapiens" /mol_type="genomic DNA" /isolate="CBRC7TM_97" /db_xref="taxon:9606" /chromosome="11" 201..1190 /codon_start=1 /evidence="not experimental" /product="seven transmembrane helix receptor" /protein_id="BAC05780.1" /db_xref="GI:21928365" /translation="MSSSTLGHNMGSPNTHVDVPSVFFLLGIFGLSQFHLSLPVCGGL GTATVGNITIVVWTEPVLHPKPVFLCLMLSTIDLAASVSTPKLAIFWCGAGHI SASACLAQWFFTHAFOMBESTVLLAMAPDYVAICPLRYATILDTIIAHIGVAIV RGLMLPCPFLIGRLNFCQSHVILHTYCEHMAVVKLAGDTRPNRVGLTAAALVIG VDLFCIGLSYALSAQNVLELSHEARSKALCTCGSHVCVILLISYTPALFSEFTTRFGH HVPVHIHILLANVILLPPALNPVTVGVTKQIRKRVVRVFGSQGMGIKASE"	321 GTGTGTGGCTTAGGCACAGCCCAAAATTTGTGGCAATATAAATACTATTCTGGTTGTGTGGCC 380
AB065534 AB065534.1 GI:21928364 SOURCE Homo sapiens (human) ORGANISM Homo sapiens Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo. 1 Suwa, M., Sato, T., Okouchi, I., Arita, M., Futami, K., Matsumoto, S., Tsutsumi, S., Aburatani, H., Asai, K. and Akiyama, Y. Genome-wide discovery and analysis of human seven transmembrane helix receptor genes Unpublished 2 (bases 1 to 1390) Suwa, M. Direct Submission Submitted (11-JUL-2001) Makiko Suwa, Computational Biology Research Center (CBRC), National Institute of Advanced Industrial Science and Technology (AIST); 2-41-6 Aomi Koto-ku, Tokyo 135-0064, Japan (E-mail: m-suwa@aist.go.jp; URL: http://www.cbrc.jp/; Tel: 81-3-3599-8080, Fax: 81-3-3599-8081) This sequence is a seven transmembrane helix receptor candidate predicted from the whole human genome sequences using our automated system that contains programs of gene finding (GeneDecoder), sequence search, motif-domain assignment and transmembrane helix prediction. And the sequence is submitted by the collaborative project between [Computational Biology Research Center (CBRC), National Institute of Advanced Industrial Science and Technology (AIST)] and [Genome Science Division, Research Center for Advanced Science and Technology (RCAT), University of Tokyo]. Location/Qualifiers 1..1390 /organism="Homo sapiens" /mol_type="genomic DNA" /isolate="CBRC7TM_97" /db_xref="taxon:9606" /chromosome="11" 201..1190 /codon_start=1 /evidence="not experimental" /product="seven transmembrane helix receptor" /protein_id="BAC05780.1" /db_xref="GI:21928365" /translation="MSSSTLGHNMGSPNTHVDVPSVFFLLGIFGLSQFHLSLPVCGGL GTATVGNITIVVWTEPVLHPKPVFLCLMLSTIDLAASVSTPKLAIFWCGAGHI SASACLAQWFFTHAFOMBESTVLLAMAPDYVAICPLRYATILDTIIAHIGVAIV RGLMLPCPFLIGRLNFCQSHVILHTYCEHMAVVKLAGDTRPNRVGLTAAALVIG VDLFCIGLSYALSAQNVLELSHEARSKALCTCGSHVCVILLISYTPALFSEFTTRFGH HVPVHIHILLANVILLPPALNPVTVGVTKQIRKRVVRVFGSQGMGIKASE"	61 ThrGluProValLeuHisLysProValTyrLeuPheLeuCysMetLeuSerThrIleAsp 80 381 ACTGAACAGTCTTGGCACAAGCCCTGTGTACCTTTTCTGTGCACTGCTCTCAACCATCGAC 440 81 LeuAlaAlaSerValSerThrValProLysLeuLeuAlaIlePheTrpCysGlyAlaGly 100 441 TTGGCTGCTCTGTCTCCACAGTTCCTCAAGCTACTGGCTATCTTCTGTGTGGAGCCGGA 500 101 HisLeuSerAlaSerAlaCysLeuAlaHisMetPhePheIleHisAlaPheCysMetMet 120 501 CATATATCTGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 560 121 GluSerThrValLeuAlaMetAlaPheAspArgTyrValAlaIleCysHisProLeu 140 561 GAGTCCACTGTGTACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 620 141 ArgTyrAlaThrIleLeuThrAspThrIleLeuAlaHisIleGlyValAlaAlaValVal 160 621 CGCTATGCCCAATCTCTCACTGACACCATCATTTGCCACATAGGGGTGGCAGCTGTAGTG 680 161 ArgGlySerLeuLeuMetLeuProCysProPhePheIleGlyArgLeuAsnPheCysGln 180 681 CGAGGCTCCCTGCTCATGCTCCCTGCTCCCTGCTCCCTGCTCCCTGCTCCCTGCTCCCTGCT 740 181 SerHisValIleLeuHisThrTyrCysGluHisMetAlaValValLysLeuAlaCysGly 200 741 AGCCATGTGATCTTACACAGTCTGTGAGCACATGGTGTGGTGAAGCTGGCTGTGGA 800 201 AspThrArgProAsnArgValTyrGlyLeuThrAlaAlaLeuLeuValIleGlyValAsp 220 801 GACACCGGCTTAACCGTGTGTGTGGCTGACAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 860 221 LeuPheCysIleGlyLeuSerTyrAlaLeuIleAlaGlnAlaValLeuAsnSerSer 240 861 TTGTTTTGCAATGGTCT 920 241 HisGluAlaAlaArgSerLysAlaLeuGlyThrCysGlySerHisValCysValIleLeu 260 921 CATGAAGCTCGTCCAGGCCCTAGGACCTGTGTGTTCCATGCTGTGTGCTGCTGCTGCTGCT 980 261 SerTyrThrProAlaLeuPheSerPhePheThrHisArgPheGlyHisHisValProVal 280 981 TCTTATACACAGCCCT 1040 281 HisIleHisIleLeuLeuAlaAsnValTyrLeuLeuLeuProProAlaLeuAsnProVal 300 1041 CATATTCACATCTTTTGGCCCAATGTTTATCTGCTTTTGGCCACCTGCTCTTAACTCTGTG 1100 301 ValTyrGlyValLysThrLysGlnIleArgLysArgValValArgValPheGlnSerGly 320 1101 GTATATGGAGTTAAGCCAAACAGATCCGTAAAGAGTTGTTCAGGGGTGTTTCAAAGTGGG 1160 321 GlnGlyMetGlyIleLysAlaSerGlu 329 1161 CAGGGAATGGGCATCAAGGCATCTGAG 1187
RESULT 8 AC026090/c LOCUS DEFINITION Homo sapiens chromosome 11 clone RP11-658K18, WORKING DRAFT SEQUENCE, 19 unordered pieces. AC026090 AC026090.5 GI:9958310 HTG; HTGS_PHASE1; HTGS_DRAFT. KEYWORDS SOURCE Homo sapiens (human) ORGANISM Homo sapiens Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo. 1 (bases 1 to 17207) Waterston, R.H. The sequence of Homo sapiens clone	

JOURNAL
REFERENCE
AUTHORS
TITLE
JOURNAL

Unpublished
2 (bases 1 to 172027)
Waterston,R.H.
Direct Submission
Submitted (19-MAR-2000) Genome Sequencing Center, Washington
University School of Medicine, 4444 Forest Park Parkway, St. Louis,
MO 63108, USA
On Sep 1, 2000 this sequence version replaced gi:9838305.

COMMENT

----- Genome Center -----
Center: Washington University Genome Sequencing Center
Center code: WUGSC
Web site: <http://genome.wustl.edu/gsc/index.shtml>
----- Project Information -----
Center project name: H.NH0658K18
----- Summary Statistics -----
Sequencing vector: M13; 100%
Sequencing vector: plasmid; 0%
Chemistry: Dye-terminator Big Dye; 0% of reads
Assembly program: Phrap; version 0.990319
Consensus quality: 16096 bases at least Q40
Consensus quality: 163968 bases at least Q30
Consensus quality: 165626 bases at least Q20
Insert size: 188000; agarose-fp
Insert size: 171387; sum-of-contigs
Quality coverage: 4.33 in Q20 bases; agarose-fp
Quality coverage: 4.82 in Q20 bases; sum-of-contigs

* NOTE: This is a 'working draft' sequence. It currently
* consists of 19 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.

1 1526: contig of 1526 bp in length
1527 1626: gap of unknown length
1627 2975: contig of 1349 bp in length
2976 3075: gap of unknown length
3076 5073: contig of 1998 bp in length
5074 5173: gap of unknown length
5174 9317: contig of 4144 bp in length
9318 9417: gap of unknown length
9418 13195: contig of 3778 bp in length
13196 13295: gap of unknown length
13296 16482: contig of 3187 bp in length
16483 21965: gap of unknown length
21966 22065: gap of unknown length
22066 33167: contig of 1102 bp in length
33168 33267: gap of unknown length
33268 43162: contig of 9895 bp in length
43163 43262: gap of unknown length
43263 53157: contig of 9895 bp in length
53158 53257: gap of unknown length
53258 61965: contig of 8708 bp in length
61966 62065: gap of unknown length
62066 70847: contig of 8782 bp in length
70848 84297: gap of unknown length
84298 84397: contig of 13350 bp in length
84398 100019: contig of 15622 bp in length
100020 100119: gap of unknown length
100120 114785: contig of 14666 bp in length
114786 114885: gap of unknown length
114886 131872: contig of 16987 bp in length
131873 131972: gap of unknown length
131973 147618: contig of 15646 bp in length
147619 171498: gap of unknown length
171499 171598: gap of unknown length

FEATURES
source

* 171599 172027: contig of 429 bp in length.
Location/Qualifiers
1..172027
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/chromosome="11"
/clone="RP11-658K18"
1..1526
/note="assembly_name:Contig14"
1627..2975
/note="assembly_name:Contig17"
3076..5073
/note="assembly_name:Contig18"
5174..9317
/note="assembly_name:Contig19"
9418..13195
/note="assembly_name:Contig20"
13296..16482
/note="assembly_name:Contig21"
16583..21965
/note="assembly_name:Contig22"
22066..33167
/note="assembly_name:Contig23"
33268..43162
/note="assembly_name:Contig24"
43263..53157
/note="assembly_name:Contig25"
53258..61965
/note="assembly_name:Contig26"
62066..70847
/note="assembly_name:Contig27"
clone_end:T7
vector_side:right
70948..84297
/note="assembly_name:Contig28"
84398..100019
/note="assembly_name:Contig29"
100120..114785
/note="assembly_name:Contig30"
114886..131872
/note="assembly_name:Contig31"
131973..147618
/note="assembly_name:Contig32"
147719..171498
/note="assembly_name:Contig33"
171599..172027
/note="assembly_name:Contig16"

ORIGIN

Alignment Scores:
Pred. No.: 5.29e-149 Length: 172027
Score: 1691.00 Matches: 325
Percent Similarity: 99.09% Conservative: 1
Best Local Similarity: 98.78% Mismatches: 3
Query Match: 98.43% Indels: 0
DB: 2 Gaps: 0

US-10-081-775-2 (1-329) x AC026090 (1-172027)

Qy 1 MetSerThrLeuGlyHisAsnMetGluSerProHisHisThrAspValAspProSer 20
Db 89467 ATGTCCAGCACTCTTGGCCCAACATGGAATCTCTTAATCACACTGATGTTGACCTTCT 89408
Qy 21 ValPhePheLeuLeuGlyIleProGlyLeuGluGlnPheHisLeuTrpLeuSerLeuPro 40
Db 89407 GTCTTCTCTCTCTGGCCATCCAGGCTCGAACAATTCATTGTGGCTCTCACTCCCT 89348
Qy 41 ValCysGlyLeuGlyThrAlaThrIleValGlyAsnIleThrIleLeuValValAla 60
Db 89347 GTGTGTGGCTTAGGCACAGCCACAAATTTGTGGCAATATACTATTCTGTGTTGTTGCC 89288
Qy 61 ThrGluProValLeuHisLysProValTyrLeuPheLeuCysMetLeuSerThrIleAsp 80

JOURNAL

Submitted (10-MAY-2002) Whitehead Institute/MIT Center for Genome Research, 320 Charles Street, Cambridge, MA 02141, USA

REFERENCE

5 (bases 1 to 176597)

AUTHORS

Birren, B., Linton, L., Nusbaum, C., Lander, E., Ali, A., Allen, N., Anderson, S., Barna, N., Bastien, V., Bloom, T., Boguslavskiy, L., Boukhgalter, B., Brown, A., Camarata, J., Campopiano, A., Chang, J., Chazaro, B., Choepel, Y., Colangelo, M., Collins, S., Collamore, A., Cook, A., Cooke, P., DeArellano, K., Dewar, K., Diaz, J. S., Dodge, S., Fero, S., Ferreira, P., FitzHugh, W., Gage, D., Galagan, J., Gardyna, S., Ginde, S., Gord, S., Goyette, M., Graham, L., Grand-Pierre, N., Hagos, B., Horton, L., Hulme, W., Iliev, I., Johnson, R., Jones, C., Kamat, A., Karatas, A., Kells, C., Lakocque, K., Lamazares, R., Landers, T., Lehoczy, J., Levine, R., Lindblad-Toh, K., Liu, G., MacLean, C., Macdonald, P., Major, J., Marquis, N., Matthews, C., McCarthy, M., McEwan, P., McKernan, K., Meldrum, J., Meneus, L., Mihova, T., Mlenga, V., Murphy, T., Naylor, J., Nguyen, C., Nicol, R., Norbu, C., Norman, C. H., O'Connor, T., O'Donnell, P., O'Neill, D., Oliver, J., Peterson, K., Phunkhang, P., Pierre, N., Pollara, V., Raymond, C., Retta, R., Rieback, M., Riley, R., Rise, C., Rogov, P., Roman, J., Rosetti, M., Roy, A., Santos, R., Schauer, S., Schupbach, R., Seaman, S., Severy, P., Spencer, B., Stange-Thomann, N., Stojanovic, N., Strauss, N., Subramanian, A., Talamas, J., Tesfaye, S., Theodore, J., Topham, K., Travers, M., Travis, N., Trigglio, J., Vassiliev, H., Viel, R., Vo, A., Wilson, B., Wu, X., Wyman, D., Ye, W. J., Young, G., Zainoun, J., Zembek, L., Zimmer, A. and Zody, M.

TITLE

Direct Submission

JOURNAL

Submitted (13-MAY-2002) Whitehead Institute/MIT Center for Genome Research, 320 Charles Street, Cambridge, MA 02141, USA

COMMENT

On May 10, 2002 this sequence version replaced gi:20177753.

All repeats were identified using RepeatMasker:

Smit, A.F.A. & Green, P. (1996-1997)

<http://ftp.genome.washington.edu/RM/RepeatMasker.html>

----- Genome Center

Center: Whitehead Institute/ MIT Center for Genome Research

Center code: WIBR

Web site: <http://www-seq.wi.mit.edu>

Contact: sequence_submissions@genome.wi.mit.edu

----- Project Information

Center project name: L26176

Center clone name: 44_D_14

FEATURES

source

Location/Qualifiers

1..176597

/organism="Homo sapiens"

/mol_type="genomic DNA"

/db_xref="taxon:9606"

/chromosome="11"

/map="11"

/clone="RP11-44D14"

/clone_lib="RPC1-11 Human Male BAC"

775..802

/rpt_family=" (TTCC)n"

1179..1224

/rpt_family=" (CA)n"

1709..2076

/rpt_family="Charlieb"

2893..2965

/rpt_family="CA-rich"

4667..4718

/rpt_family="GA-rich"

6660..6815

/rpt_family="L2"

complement (6896..7766)

/rpt_family="L1M2"

complement (8463..8683)

/rpt_family="L1M4"

8684..8720

/rpt_family=" (TTTA)n"

complement (8721..9405)

/rpt_family="L1P4"

complement (9406..9469)

/rpt_family="L1M4"

complement (9486..9561)

repeat_region /rpt_family="FLAM" 10351..10432
repeat_region /rpt_family="HAL1" 10833..10896
repeat_region /rpt_family=" (TA)n" complement (12312..12520)
repeat_region /rpt_family="MIR" 13585
repeat_region /rpt_family="MER77" complement (14764..15303)
repeat_region /rpt_family="L1MBS" complement (15304..15618)
repeat_region /rpt_family="AluY" complement (15619..16311)
repeat_region /rpt_family="L1MBS" complement (16312..16610)
repeat_region /rpt_family="AluSc" complement (16611..16666)
repeat_region /rpt_family="L1MBS" complement (16677..16816)
repeat_region /rpt_family="L1MA6" 16815..19052
repeat_region /rpt_family="L1PA5" complement (19063..19103)
repeat_region /rpt_family="L1MA6" complement (19102..20472)
repeat_region /rpt_family="L1PA16" complement (20467..20806)
repeat_region /rpt_family="L1MBS" 20806..22902
repeat_region /rpt_family="L1MBS" complement (22888..22923)
repeat_region /rpt_family="L1MA7" complement (22937..23722)
repeat_region /rpt_family="L1M4" complement (23871..24202)
repeat_region /rpt_family="L1M4" 25092..25416

Alignment Scores:

Pred. No.: 5,46e-149 Length: 176597
Score: 1691.00 Matches: 325
Percent Similarity: 99.09% Conservative: 1
Best Local Similarity: 98.78% Mismatches: 3
Query Match: 98.43% Indels: 0
DB: 9 Gaps: 0

US-10-081-775-2 (1-329) x AC116156 (1-176597)

Qy 1 MetSerThrLeuGlyHisAsnMetGluSerProHisHisThrAspValAspProSer 20
Db 159269 ATGTCCAGCACTCTTGGCCACACATGGAATCTCTTAATCACACTGATGTTGACCTTCT 159210
Qy 21 ValPhePheLeuLeuGlyIleProGlyLeuGluGlnPheHisLeuTrpLeuSerLeuPro 40
Db 159209 GTCTTCTCTCTCCCTGGGCATCCCAAGGCTCGAACAAATTTATTGTGGCTCTCACTCCT 159150
Qy 41 ValCysGlyLeuGlyThrAlaThrIleValGlyAsnIleThrIleLeuValValValAla 60
Db 159149 GTGTGTGGCTTGGCCAGCCACACAAATTTGTGGCAATATACTATTCTGGTGTGTGGCC 159090
Qy 61 ThrGluProValLeuHisLysProValTyrLeuPheLeuCysMetLeuSerThrIleAsp 80
Db 159089 ACTGAACCACTGTGCACAGCCTGTGTACTTTTCTGTGATGCTCTCAACCATCGAC 159030
Qy 81 LeuAlaAlaSerValSerThrValProLysLeuLeuAlaIlePheTrpCysGlyAlaGly 100
Db 159029 TTGGTGCCTCTGTCTCCACAGTTCCCAAGCTACTTGCTACTTCTCTGTGGAGCCGA 158970
Qy 101 HisIleSerAlaSerAlaCysLeuAlaHisMetPhePheIleHisAlaPheCysMetMet 120
Db 158969 CATATATCGCCTCTGCCTGCTGGCCACAGATGTTCTTCAATGCTTCTGATGATG 158910

Qy 121 GluSerThrValLeuLeuAlaMetAlaPheAspArgTyrValAlaIleCysHisProLeu 140
 Db 158909 GAGTCCACTGTGCTACTGGCCATGCGCTTTGATCGTACGTGGCCATCTGCACCCACTC 158850
 Qy 141 ArgTyrAlaThrIleLeuThrAspThrIleIleAlaHisIleGlyValAlaAlaValVal 160
 Db 158849 CGCTATGACCAATCTCTACACCATCATTTGCCACATAGGGGTGGCAGCTGTAGTG 158790
 Qy 161 ArgGlySerLeuLeuMetLeuProCysProPhePheIleGlyArgLeuAsnPheCysGln 180
 Db 158789 CGAGCTCCCTGCTCATGCTCCCATGTCCCTTCTTATTTGGCGCTTTGAACCTTCTGCCAA 158730
 Qy 181 SerHisValIleLeuHisThrTyrCysGluHisMetAlaValValLeuAlaCysGly 200
 Db 158729 AGCCATGTGATCTTACACACGTACTGTGAGCACATGGCTGTGGTGAAGCTGGCCCTGTGGA 158670
 Qy 201 AspThrArgProAsnArgValTyrGlyLeuThrAlaAlaLeuValIleGlyValAsp 220
 Db 158669 GACACAGCGCTTACCGTGTGTATGGGTGACAGCTGACGTGTGCTATTTGGGTGAC 158610
 Qy 221 LeuPheCysIleGlyLeuSerTyrAlaLeuIleAlaGlnAlaValLeuArgLeuSerSer 240
 Db 158609 TTGTTTTGCATTGGTCTCTCTATGCCCTAAGTGACAAAGCTGTCTTCCGCTCTCATCC 158550
 Qy 241 HisGluAlaArgSerIleAlaLeuGlyThrCysGlySerHisValCysValIleLeuIle 260
 Db 158549 CATGAAGCTCGTCCAAAGCCCTTAGGACCTGTGTGTTCCCATGTCTGTCTCATCTCATC 158490
 Qy 261 SerTyrThrProAlaLeuPheSerPhePheThrHisArgPheGlyHisValProVal 280
 Db 158489 TCTTATACACCGCCCTCTCTCTCTTTTACACACCGCTTTGGCCATCAGCTTCCAGTC 158430
 Qy 281 HisIleHisIleLeuLeuAlaAsnValTyrLeuLeuLeuProAlaLeuAsnProVal 300
 Db 158429 CATATTCAATCTTTGGCCAAATGTTATCTGTCTTTGGCCACCTGCTCTTAATCTGTG 158370
 Qy 301 ValTyrGlyValIleThrIleGlnIleArgIleArgValValArgValPheGlnSerGly 320
 Db 158369 GTATATGAGGTAAAGACCAACAGATCCGTAAGAGGTGTTCAGGGTGTTCAAAGTGGG 158310
 Qy 321 GlnGlyMetGlyIleLeuAlaSerGlu 329
 Db 158309 CAGGGAATGGGCATCAAGGCATCTCAG 158283

RESULT 10
 AX241675
 LOCUS AX241675 963 bp DNA linear PAT 26-SEP-2001
 DEFINITION Sequence 423 from Patent WO0127158.
 ACCESSION AX241675
 VERSION AX241675.1 GI:15798550
 KEYWORDS synthetic construct
 SOURCE synthetic construct
 ORGANISM artificial sequences.
 REFERENCE 1
 AUTHORS Bellenson,J., Smith,D., Lancet,D., Glusman,G., Fuchs,T. and Yanai,I.
 TITLE Olfactory receptor sequences
 JOURNAL Patent: WO 0127158-A 423 19-APR-2001;
 DIGISCENTS (US); YEDA RESEARCH AND DEVELOPMENT COMPANY, LTD. (IL)
 FEATURES
 Location/Qualifiers
 source 1..963
 /organism="synthetic construct"
 /mol_type="unassigned DNA"
 /db_xref="taxon:32630"
 /note="(H38g272 nucleotide)"
 ORIGIN
 Alignment Scores:
 Pred. No.: 1.26e-147 Length: 963
 Score: 1649.00 Matches: 317
 Percent Similarity: 99.07% Conservative: 1
 Best Local Similarity: 98.75% Mismatches: 3

Query Match: 95.98% Indels: 0
 DB: 6 Gaps: 0
 US-10-081-775-2 (1-329) x AX241675 (1-963)
 Qy 9 MetGluSerProHisHisThrAspValAspProSerValPhePheLeuLeuGlyIlePro 28
 Db 1 ATGGAATCTCTAATCACACTGATGTTGACCTTCTGTCTTCTTCTCTCTCTCTCTCTCA 60
 Qy 29 GlyLeuGluGlnPheHisLeuThrIleLeuSerLeuProValCysGlyLeuGlyThrAlaThr 48
 Db 61 GGTCTGGAACAATTTCAATTTGTGCTCTCACTCCCTGTGTGTGGCTTAGGCACAGCCACA 120
 Qy 49 IleValGlyAenIleThrIleLeuValValAlaThrGluProValLeuHisIlePro 68
 Db 121 ATTTGGGCAATAATAACTATTCTGTGTGTGTGGCTGACCTGACCACTGTTGCACAGGCT 180
 Qy 69 ValTyrLeuPheLeuCysMetLeuSerThrIleAspLeuAlaAlaSerValSerThrVal 88
 Db 181 GTGTACCTTTTCTGTGTCATGCTCTCAACCATCGACTTGGCTGCTCTGTCTTCCACAGTT 240
 Qy 89 ProLysLeuLeuAlaIlePheThrCysGlyAlaGlyHisIleSerAlaSerAlaCysLeu 108
 Db 241 CCCAAGCTTACTGGCTATCTTCTGTGTGGAGCCGACATATATCTGCTCTGCTCTGCTCTG 300
 Qy 109 AlaHisMetPhePheIleHisAlaPheCysMetMetGluSerThrValLeuLeuAlaMet 128
 Db 301 GCACAGATGTTCTTCAATTCATGCTCTGTCATGATGGAGTCCACCTGCTACTGCTACTG 360
 Qy 129 AlaPheAspArgTyrValAlaIleCysHisProLeuArgTyrAlaThrIleLeuThrAsp 148
 Db 361 GCCTTTGATCGTACGTGGCCATCTGCCACCATCCGCTATGCCACCAATCTCTCACTGAC 420
 Qy 149 ThrIleIleAlaHisIleGlyValAlaValAlaValArgGlySerLeuLeuMetLeuPro 168
 Db 421 ACCATCATGTGCCACATAGGGGTGGCAGCTGTAGTGGAGGCTCCCTGCTCTGCTCTGCT 480
 Qy 169 CysProPhePheIleGlyArgLeuAsnPheCysGlnSerHisValIleLeuHisThrTyr 188
 Db 481 TGTCTCTCTCTTATTTGGCGGTTTGAACCTTCTGCCAAAGCCATGTGATCTCTACACGTA 540
 Qy 189 CysGluHisMetAlaValValLysLeuAlaCysGlyAspThrArgProAsnArgValTyr 208
 Db 541 TGTGAGCACATGGCTGTGTGAAGCTGGCTGTGGAGACACCCAGGCTTAACCGTGTGTAT 600
 Qy 209 GlyLeuThrAlaAlaLeuValIleGlyValAspLeuPheCysIleGlyLeuSerTyr 228
 Db 601 GGGCTGACAGCTGCACTGTGTGCTCACTGGGGTTGACTTGTGTTTGTGCTTCTCTCTAT 560
 Qy 229 AlaLeuIleAlaGlnAlaValLeuArgLeuSerSerHisGluAlaArgSerLysAlaLeu 248
 Db 661 GCCCTAAGTGACCAAGCTGCTCTTCCCTCTCATCCCAAGCTCGGTCCAAGGCCCTA 720
 Qy 249 GlyThrCysGlySerHisValCysValIleLeuIleSerTyrThrProAlaLeuPheSer 268
 Db 721 GGGACCTGTGGTTCCTCATGCTGTGTGCTCATCTCTTATACACACCGCTCTCTCTCC 780
 Qy 269 PhePheThrHisArgPheGlyHisValProValHisIleHisIleLeuLeuAlaAsn 288
 Db 781 TTTTATTACACACCGCTTTGGCCATCAGCTTCCAGTCCATATTTCACATCTTTTGGCAAT 840
 Qy 289 ValTyrLeuLeuLeuProProAlaLeuAsnProValValTyrGlyValLysThrLysGln 308
 Db 841 GTTTATCTGCTTTTGGCCACTGCTCTTAATCTGTGTATATGGAGTTAAGACCAACAG 900
 Qy 309 IleArgLysArgValValArgValPheGlnSerGlyGlnGlyMetGlyIleLysAlaSer 328
 Db 901 ATCCGTAAAGAGATTGTGAGGGTGTTCAAAGTGGCAGGGAATGGGCATCAAGGATCT 960
 Qy 329 Glu 329
 Db 961 GAG 963

```

RESULT 11
AX318261
LOCUS AX318261 966 bp DNA linear PAT 14-DEC-2001
DEFINITION Sequence 42 from Patent WO0190359.
ACCESSION AX318261
VERSION AX318261.1 GI:17900921
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
1. Patterson, C., Tribouley, C.M., Yao, M.G., Griffin, J.A., Thornton, M.,
Lu, Y., Kallick, D.A., Gandhi, A.R. and Au-Young, J.
G-protein coupled receptors
Patent: WO 0190359-A 42 29-NOV-2001;
Incyte Genomics, Inc. (US)
FEATURES
source
1..966
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"
/note="Incyte ID No: 7476077CBI"
ORIGIN
Alignment Scores:
Pred. No.: 1,27e-147 Length: 966
Score: 1649.00 Matches: 317
Percent Similarity: 99.07% Conservative: 1
Best Local Similarity: 98.75% Mismatches: 3
Query Match: 95.98% Indels: 0
DB: 6 Gaps: 0
US-10-081-775-2 (1-329) x AX318261 (1-966)
QY 9 MetGluSerProHisHisThrAspValAspProSerValPhePheLeuLeuGlyIlePro 28
Db 1 ATGGAATCTCTAATACACATGATGTGACCTTCTGCTTCTCTCTCTGCGCATCCCA 60
QY 29 GlyLeuGluGlnPheHisLeuThrPheSerLeuProValCysGlyLeuGlyThrAlaThr 48
Db 61 GGTCTGGAACAATATTCATTTGTGGCTCTCACTCCCTGTGTGGCTAGGCACACCA 120
QY 49 IleValGlyAsnIleThrIleLeuValValAlaThrGluProValLeuHisIysPro 68
Db 121 ATGTGGGCAATATAACATATCTGTTGTGTGGCTCACTCCCTGTGTGGCTAGGCAC 180
QY 69 ValTyrLeuPheLeuCysMetLeuSerThrIleAspLeuAlaAlaSerValSerThrVal 88
Db 181 GTGTACCTTTTCTGTGCATGCTCTCAACCATCGACTTGGCTGCTGTCTCCACAGTT 240
QY 89 ProIysLeuLeuAlaIlePheThrPheCysGlyAlaGlyHisIleSerAlaSerAlaCysLeu 108
Db 241 CCCAAGCTACTGGCTATCTTCTGTGTGGAGCCGACATATATCTGCTCTGCTGCTG 300
QY 109 AlaHisMetPhePheIleHisAlaPheCysMetMetGluSerThrValLeuLeuAlaMet 128
Db 301 GCACAGATGTTCTTCATTCAGCTCTTGTGATGAGTCCACTGTCTGCTGCTGCTG 360
QY 129 AlaPheAspArgTyrValAlaIleCysHisProLeuArgTyrAlaThrIleLeuThrAsp 148
Db 361 GCCTTTGATCGTACGTGGCCATCTGCCACCATCCGCTATGCCATCTCTCCTGAC 420
QY 149 ThrIleIleAlaHisIleGlyValAlaAlaValValArgGlySerLeuLeuMetLeuPro 168
Db 421 ACCATCATTTGCCACATAGGGGTGGCAGCTGTAGTGGAGGCTCCCTGCTCATGCTCCA 480
QY 169 CysProPhePheIleGlyArgLeuAsnPheCysGlnSerHisValIleLeuHisThrTyr 188
Db 481 TGTCCCTTCCTATTGGCGGTTTGAACCTCTGCCAAAGCCCATGTGATCTCTACACAG 540
QY 189 CysGluHisMetAlaValValLysLeuAlaCysGlyAspThrArgProAsnArgValTyr 208

```

```

Db 541 TGTGAGCAGCATGGCTGTGGTGAAGCTGCGCTGTGGAGACACAGCGCCTAACCGTGTGTAT 600
QY 209 GlyLeuThrAlaAlaLeuLeuValIleGlyValAspLeuPheCysIleGlyLeuSerTyr 228
Db 601 GGGCTGACAGCTGCACCTGTGGTCATTTGGGTGACCTTGTGCTTGTGCTCTCTCTAT 660
QY 229 AlaLeuIleAlaGlnAlaValLeuArgLeuSerSerHisGluAlaArgSerLysAlaLeu 248
Db 661 GCCCTAAGTGACAAAGCTGCTCTTGGCTCTCATCCCATGAAGCTCGGTCCAGGCCCTA 720
QY 249 GlyThrCysGlySerHisValCysValIleLeuIleSerTyrThrProAlaLeuPheSer 268
Db 721 GGGACCTGTGGTTCCTCATCTGTGTCTCATCTCTCTTATACACAGCCTCTCTCTCC 780
QY 269 PhePheThrHisArgPheGlyHisValProValHisIleHisIleLeuLeuAlaAsn 288
Db 781 TTTTATACACACCGCTTTGGCCATCAGCTTCAGTCCATATTCACATCTCTTTTGGCCA 840
QY 289 ValTyrLeuLeuLeuProProAlaLeuAsnProValValTyrGlyValLysThrIysGln 308
Db 841 GTTATCTGCTTTTGGCCACCTGCTCTTAATCTGTGTATATGAGTTAAGACCAACAG 900
QY 309 IleArgLysArgValValArgValPheGlnSerGlyGlnGlyMetGlyIleLysAlaSer 328
Db 901 ATCCGTAAAGAGTTGTTCAGGGTGTTCAAAGTGGGAGGGAATGGGCAATCAAGCACT 960
QY 329 Glu 329
Db 961 GAG 963
RESULT 12
AX448757
LOCUS AX448757 966 bp DNA linear PAT 03-JUL-2002
DEFINITION Sequence 423 from Patent WO0224726.
ACCESSION AX448757
VERSION AX448757.1 GI:21697655
KEYWORDS Homo sapiens (human)
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
1. Veithen, A.
Olfactory and pheromones g-protein coupled receptors
Patent: WO 0224726-A 423 28-MAR-2002;
ChemCom S.A. (BE)
FEATURES
source
1..966
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"
1..966
/note="unnamed protein product"
CDS
1..966
/codon_start=1
/protein_id="CAD37694.1"
/db_xref="GI:21697656"
/translat="MESPNTDVPVSFVFLGIPGLEQFHLWLSLPVCGLTATVGN
ITIVVATPEVLHKPVYLCMLSTIDLAASVSTVKLLAI FWCAGHISASACLAQ
MFTIAPCMESVLLAMAFDRVAICHPLRYATILTDITIAHIGVAUVRGSLMLP
CPFLIQLNFCQSHVILHTCEHMAVVVVLACGTRPNRYVGLTAALIVIGVDLFCIGL
SYALSQAQVLRLSSHEARSKALGTGSHVCVILISYTPFLSFPTFRFHGHVPHIHI
LLANVYLLLPALNPVYVYVKTKQIRKRVVRVFSQGGMGKASE"
ORIGIN
Alignment Scores:
Pred. No.: 1,27e-147 Length: 966
Score: 1649.00 Matches: 317
Percent Similarity: 99.07% Conservative: 1
Best Local Similarity: 98.75% Mismatches: 3
Query Match: 95.98% Indels: 0
DB: 6 Gaps: 0

```


Kamat, A., Karatas, A., Kells, C., Labocque, K., Lamazares, R., Landers, T., Lehoczy, J., Levine, R., Lindblad-Toh, K., Liu, G., MacLean, C., MacDonald, P., Major, J., Marquis, N., Matthews, C., McCarthy, M., McEwan, P., McKernan, K., Meldrum, J., Meneus, L., Mihova, T., Mlenga, V., Murphy, T., Naylor, J., Nguyen, C., Nicol, R., Norbu, C., Norman, C.H., O'Connor, T., O'Donnell, P., O'Neill, D., Oliver, J., Peterson, K., Phunkhang, P., Pierre, N., Pollara, V., Raymond, C., Retta, R., Rieback, M., Riley, R., Rise, C., Rogov, P., Roman, J., Rosetti, M., Roy, A., Santos, R., Schauer, S., Schupbach, R., Seanan, S., Severy, P., Spencer, B., Stange-Thomann, N., Stojanovic, N., Strauss, N., Subramanian, A., Talamas, J., Tesfaye, S., Theodores, J., Topham, K., Travers, M., Travis, N., Trigilio, J., Vassiliev, H., Viel, R., Vo, A., Wilson, B., Wu, X., Wyman, D., Ye, W.J., Young, G., Zainoun, J., Zembek, L., Zimmer, A. and Zody, M.

Direct Submission

Submitted (14-MAY-2002) Whitehead Institute/MIT Center for Genome Research, 320 Charles Street, Cambridge, MA 02141, USA

On May 14, 2002 this sequence version replaced gi:20429487.

All repeats were identified using RepeatMasker:

Smit, A.F.A. & Green, P. (1996-1997)

<http://ftp.genome.washington.edu/RM/RepeatMasker.html>

----- Genome Center

Center: Whitehead Institute/ MIT Center for Genome Research

Center code: WIBR

Web site: <http://www-seq.wi.mit.edu>

Contact: sequence_submissions@genome.wi.mit.edu

----- Project Information

Center project name: L24591

Center clone name: 2504_M_7

FEATURES

source	Location/Qualifiers
1..202761	/organism="Homo sapiens"
	/mol_type="genomic DNA"
	/db_xref="taxon:9606"
	/chromosome="11"
	/map="11"
	/clone="CTD-2504M7"
	/clone_lib="CITD2 Human BAC"
repeat_region	243..263
repeat_region	336..441
repeat_region	/rpt_family="AT_rich"
repeat_region	/rpt_family="L3"
repeat_region	complement(445..610)
repeat_region	/rpt_family="L1PA13"
repeat_region	complement(628..765)
repeat_region	/rpt_family="L3"
repeat_region	complement(2274..3650)
repeat_region	/rpt_family="Tigger1"
repeat_region	3651..3995
repeat_region	/rpt_family="AluSq"
repeat_region	3995..4033
repeat_region	/rpt_family="(GA)n"
repeat_region	complement(3996..4748)
repeat_region	/rpt_family="Tigger1"
repeat_region	4749..4971
repeat_region	/rpt_family="L1M4"
repeat_region	5007..6333
repeat_region	/rpt_family="L1MA9"
repeat_region	6456..6697
repeat_region	/rpt_family="L1MB1"
repeat_region	complement(7329..7491)
repeat_region	/rpt_family="L1PA12"
repeat_region	7495..7751
repeat_region	/rpt_family="L1MB1"
repeat_region	8985..9007
repeat_region	/rpt_family="AT_rich"
repeat_region	11044..11180
repeat_region	/rpt_family="L1MB1"
repeat_region	11713..12142
repeat_region	/rpt_family="L1PA5"
repeat_region	12483..12527
repeat_region	/rpt_family="AT_rich"

repeat_region	12636..12838	repeat_region	12636..12838
repeat_region	/rpt_family="HAL1"	repeat_region	/rpt_family="HAL1"
repeat_region	complement(13044..13443)	repeat_region	complement(13044..13443)
repeat_region	/rpt_family="L1MA9"	repeat_region	/rpt_family="L1MA9"
repeat_region	complement(13639..13939)	repeat_region	complement(13639..13939)
repeat_region	/rpt_family="AluSq"	repeat_region	/rpt_family="AluSq"
repeat_region	16194..16222	repeat_region	16194..16222
repeat_region	/rpt_family="AT_rich"	repeat_region	/rpt_family="AT_rich"
repeat_region	complement(17502..17712)	repeat_region	complement(17502..17712)
repeat_region	/rpt_family="L1MA9"	repeat_region	/rpt_family="L1MA9"
repeat_region	complement(17688..18018)	repeat_region	complement(17688..18018)
repeat_region	/rpt_family="L1MA9"	repeat_region	/rpt_family="L1MA9"
repeat_region	complement(18247..18311)	repeat_region	complement(18247..18311)
repeat_region	/rpt_family="L2"	repeat_region	/rpt_family="L2"
repeat_region	complement(19166..19280)	repeat_region	complement(19166..19280)
repeat_region	/rpt_family="MIR"	repeat_region	/rpt_family="MIR"
repeat_region	19450..19637	repeat_region	19450..19637
repeat_region	/rpt_family="MIR"	repeat_region	/rpt_family="MIR"
repeat_region	21072..21116	repeat_region	21072..21116
repeat_region	/rpt_family="L1MB1"	repeat_region	/rpt_family="L1MB1"
repeat_region	complement(21146..21534)	repeat_region	complement(21146..21534)
repeat_region	/rpt_family="L1MA4"	repeat_region	/rpt_family="L1MA4"
repeat_region	21536..21774	repeat_region	21536..21774
repeat_region	/rpt_family="L1MA4"	repeat_region	/rpt_family="L1MA4"
repeat_region	22101..22162	repeat_region	22101..22162
repeat_region	/rpt_family="(TTCA)n"	repeat_region	/rpt_family="(TTCA)n"
repeat_region	23080..23388	repeat_region	23080..23388
repeat_region	/rpt_family="AluSq"	repeat_region	/rpt_family="AluSq"
repeat_region	23817..23955	repeat_region	23817..23955
repeat_region	/rpt_family="Charliel"	repeat_region	/rpt_family="Charliel"
repeat_region	complement(24081..24150)	repeat_region	complement(24081..24150)
repeat_region	/rpt_family="L1MB3"	repeat_region	/rpt_family="L1MB3"
repeat_region	complement(24151..24251)	repeat_region	complement(24151..24251)
repeat_region	/rpt_family="L1MA1"	repeat_region	/rpt_family="L1MA1"
repeat_region	24254..24829	repeat_region	24254..24829
repeat_region	/rpt_family="L1MA1"	repeat_region	/rpt_family="L1MA1"
repeat_region	complement(24841..25271)	repeat_region	complement(24841..25271)
repeat_region	/rpt_family="L1MB6"	repeat_region	/rpt_family="L1MB6"
repeat_region	25424..25503	repeat_region	25424..25503
repeat_region	/rpt_family="(TA)n"	repeat_region	/rpt_family="(TA)n"
repeat_region	25712..26331	repeat_region	25712..26331
repeat_region	/rpt_family="L1MA2"	repeat_region	/rpt_family="L1MA2"
repeat_region	26719..26754	repeat_region	26719..26754
repeat_region	/rpt_family="AT_rich"	repeat_region	/rpt_family="AT_rich"
repeat_region	27022..27136	repeat_region	27022..27136
repeat_region	/rpt_family="(TTATA)n"	repeat_region	/rpt_family="(TTATA)n"
repeat_region	28928..28957	repeat_region	28928..28957
repeat_region	/rpt_family="AT_rich"	repeat_region	/rpt_family="AT_rich"
repeat_region	29743..30367	repeat_region	29743..30367
repeat_region	/rpt_family="L1PB3"	repeat_region	/rpt_family="L1PB3"
repeat_region	complement(30373..30691)	repeat_region	complement(30373..30691)
repeat_region	/rpt_family="L1MB6"	repeat_region	/rpt_family="L1MB6"
repeat_region	complement(30926..30987)	repeat_region	complement(30926..30987)

Alignment Scores:	8.89e-145	Length:	202761
Pred. No.:	1647.50	Matches:	324
Score:	98.18%	Conservative:	0
Percent Similarity:	98.18%	Mismatches:	4
Best Local Similarity:	95.90%	Indels:	3
Query Match:	9	Gaps:	1
DB:			

US-10-081-775-2 (1-329) x AC109341 (1-202761)
QY 1 MetSerThrLeuGlyHisAsnMetGluSerProHisThrAppValAspProSer 20
Db 69631 ATGTCAGACACTTTGGCCACAAATCGGAATCCTCATCACACTGATGTGACCTTCT 69572
QY 21 -ValPhePheLeuGlyIleProGlyLeuGluInPheHisLeuTrpLeuSerLeuPr 40
Db 69571 TGTCTTCTTCTCTCTGGCATCCAGGTCGTGGACAAATTCATTGTGGCTCTCCTCCC 69512
QY 40 oValCysGlyLeuGlyThrAlaThrIleValGlyAsnIleThrIleLeuValValAl 60

```
Db 69511 TGTGTGGCTTATGGACAGCCACAAATGTGGCAATATAACTATTCGTGTGTGTGGC 69452
Qy 60 aThrGluProValLeuHisLysProValTyrLeuPheLeuCysMetLeuSerThrIleAs 80
Db 69451 CACTGAACCAAGTCTTGACAGACCTGTGTACTCTTTCTTGTCATCTCTCAACCATCGA 69392
Qy 80 pLeuAlaSerValSerThrValProLysLeuLeuAlaIlePheTyrCysGlyAlaG1 100
Db 69391 CTTGGCTGCTCTGTCTCCACAGTTCCTCCAGCTACTGGCTATCTCTGTGTGGAGCGG 69332
Qy 100 yHisIleSerAlaSerAlaCysLeuAlaHisMetPhePheIleHisAlaPheCysMetMe 120
Db 69331 ACATATATCTGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 69272
Qy 120 tGluSerThrValLeuLeuAlaMetAlaPheAspArgTyrValAlaIleCysHisProLe 140
Db 69271 GGAGTCCACTGTGCTACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 69212
Qy 140 uArgTyrAlaThrIleLeuThrAspThrIleIleAlaHisIleGlyValAlaAlaValVa 160
Db 69211 CCGCTATGCCACATCCTCACTGACACCATCATTTGCCACATGAGGGTGGCAGCTGTAGT 69152
Qy 160 lArgGlySerLeuLeuMetLeuProCysProPhePheIleGlyArgLeuAsnPheCysG1 180
Db 69151 GCGAGGCTCCCTGCTCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 69092
Qy 180 nSerHisValIleLeuHisThrTyrCysGluHisMetAlaValValLysLeuAlaCysG1 200
Db 69091 AAGCCATGTGATCTTACACACATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 69036
Qy 200 yAspThrArgProAsnArgValTyrGlyLeuThrAlaAlaLeuValIleGlyValAs 220
Db 69035 AGACACCAAGGCTAAACCGTGTGTATGGCTGACATGCTGCTGCTGCTGCTGCTGCTGCTG 68976
Qy 220 pLeuPheCysIleGlyLeuSerTyrAlaLeuIleAlaGlnAlaValLeuArgLeuSerSe 240
Db 68975 CTTGTTTTCATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 68916
Qy 240 tHisGluAlaArgSerLysAlaLeuGlyThrCysGlySerHisValCysValIleLeu1 260
Db 68915 CCATGAAGCTGGTCCAAAGGCCCTAGGACCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCT 68856
Qy 260 eSerTyrThrProAlaLeuPheSerPhePheThrHisArgPheGlyHisHisValProVa 280
Db 68855 CTCCTATACACAGCCCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 68796
Qy 280 HisIleHisIleLeuLeuAlaAsnValTyrLeuLeuLeuProProAlaLeuAsnProVa 300
Db 68795 CCATATTCATCTCTTTGGCCAAATGTTTATCTGCTTTGGCCACCTGCTCTTAATCTGT 68736
Qy 300 lValTyrGlyValLysThrLysGlnIleArgLysArgValValValArgValPheGlnSerG1 320
Db 68735 GGTATATGAGTAAAGCAACACAGATCCGTAAAGAGTGTGAGGGTGTCTTCAAAGTGG 68676
Qy 320 yClnGlyMetGlyIleLysAlaSerCln 329
Db 68675 GCAGAAATGGGCATCAAGGCATCTGAG 68648

RESULT 14
AC147586
LOCUS AC147586 181386 bp DNA linear HTG 18-DEC-2003
DEFINITION Otollemur garnettii clone CH256-553F22, WORKING DRAFT SEQUENCE.
ACCESSION AC147586
VERSION AC147586.1 GI:40018680
KEYWORDS HTG; HTGS PHASE2; HTGS DRAFT.
SOURCE Otollemur garnettii (small-eared galago)
ORGANISM Otollemur garnettii
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Strepsirhini; Galgonidae; Otollemur.
1 (bases 1 to 181386)
REFERENCE
AUTHORS Antonellis,A., Ayele,K., Benjamin,B., Blakesley,R.W.,
Bouffard,G.G., Brinkley,C., Brooks,S., Chu,G., Coleman,B.,
Coleman,H., Daki,N., Engle,J., Granite,S., Guan,X., Gupta,J.,
Haghighi,P., Han,J., Hansen,N., Ho,S.-L., Hu,P., Hurle,B.,
Idol,R., Jones,C., Karlins,E., Kim,H., Kwong,P., Laric,P.,
Larson,S., Lee-Jin,S.-Q., Legaspi,R., Maduro,Q.L., Maduro,V.B.,
Margulies,E.H., Masiello,C., Maskeri,B., McDowell,J.,
Mullikin,J.C., Paguirigan,C., Portnoy,M.E., Prasad,A., Puri,O.,
Reddi-Dugue,N., Schandler,K., Schueler,M.G., Shah,K., Sison,C.,
Stantrop,S., Thomas,J.W., Thomas,P.J., Tsipouri,V., Vogt,J.L.,
Weherby,K.D., Young,A. and Green,E.D.
NISC Comparative Sequencing Initiative
Unpublished
2 (bases 1 to 181386)
Green,E.D.
Direct Submission
Submitted (18-DEC-2003) NIH Intramural Sequencing Center, 8717
Gromont Circle, Gaithersburg, MD 20877, USA
----- Genome Center
Center: NIH Intramural Sequencing Center
Center code: NISC
Web site: http://www.nisc.nih.gov
Contact: nisc.zoo@hri.nih.gov
----- Project Information
Center project name: f11
Center clone name: 553F22
----- Summary Statistics
Sequencing vector: plasmid; n/a; 100% of reads
Chemistry: Dye-terminator Big Dye; 100% of reads
Assembly program: Phrap; version 0.990319
Consensus quality: 180937 bases at least Q40
Consensus quality: 181081 bases at least Q30
Consensus quality: 181163 bases at least Q20
Insert size: 177000; agarose-fp
Quality coverage: 13.31x in Q20 bases; agarose-fp
Quality coverage: 12.99x in Q20 bases; sum-of-contigs
-----
* NOTE: This is a 'working draft' sequence. It currently
* consists of 1 contigs. Gaps between the contigs
* are represented as runs of N. The order of the pieces
* is believed to be correct as given, however the sizes
* of the gaps between them are based on estimates that have
* provided by the submitter.
* This sequence will be replaced
* by the finished sequence as soon as it is available and
* the accession number will be preserved.
*
* 1 181386: contig of 181386 bp in length.
-----
Location/Qualifiers
1..181386
/organism="Otollemur garnettii"
/mol_type="genomic DNA"
/db_xref="taxon:30611"
/clone="CH256-553F22"
/clone_lib="CH256"
/note="BAC resource: http://bacpac.chori.org/"
1..181386
/note="assembly_fragment"
clone_end:SP6
vector_side:right"

FEATURES
source
misc_feature
1..181386
/note="assembly_fragment"
clone_end:SP6
vector_side:right"

ORIGIN
Alignment Scores:
Pred. No.: 2,6e-136 Length: 181386
Score: 1558.00 Matches: 297
Percent Similarity: 94.22% Conservative: 13
Best local Similarity: 90.27% Mismatches: 19
Query Match: 90.69% Indels: 0
DB: Gaps: 0
US-10-081-775-2 (1-329) x AC147586 (1-181386)
Qy 1 MetSerSerThrLeuGlyHisAsnMetGluSerProHisHisThrAspValAspProSer 20
Db 3117 ATATCCGGTACTCTTGGCCAGATCATGAAATCTCTTAATCACACACATGGACCTTCT 3176
```

```

QY 21 ValPhePheLeuLeuGlyLeuProGlyLeuGluGlnPheHisLeuTrpLeuSerLeuPro 40
Db ::::|||||
3177 ATTTTCTTCTCTTGGGTATTTCAGGTCTAGAAAAAATTTTCATCTGGCTCTCATCTCT 3236

QY 41 ValCysGlyLeuGlyThrAlaThrIleValGlyAsnIleThrIleLeuValValAla 60
Db ::::|||||
3237 GTGTGCTGCTGGGCACAGCCAGTTGTGGCAACATAACCATCTCTGTTGTGTGGCC 3296

QY 61 ThrGluProValLeuHisLysProValTyrLeuPheLeuCysMetLeuSerThrIleAsp 80
Db ::::|||||
3297 ACTGAGCCAGCCCTGCACAGCTGTGTACCTTTTCTCTGTCATGTGTCAACCATCGAC 3356

QY 81 LeuAlaAlaSerValSerThrValProLysLeuLeuAlaIlePheThrCysGlyAlaGly 100
Db ::::|||||
3357 TTGGCTGCTCTGTCTCCACAGTTCCCAAGCTACTGGGCATTTCTCTGGTGTGGAGCTGA 3416

QY 101 HisIleSerAlaSerAlaCysLeuAlaHisMetPheIleHisAlaPheCysMetMet 120
Db ::::|||||
3417 CATATATCTGCTCTGCTTGCCTAGCACAGATGTTCTTCATTCATGCGCTTCTGCATGATG 3476

QY 121 GluSerThrValLeuLeuAlaMetAlaPheAspArgTyrValAlaIleCysHisProLeu 140
Db ::::|||||
3477 GAGTCCATCTGCTGCTGCCATGCGCTTGTATCGCTATGTGCCATCTGCGCATCCATCT 3536

QY 141 ArgTyrAlaThrIleLeuThrAspThrIleAlaHisIleGlyValAlaValVal 160
Db ::::|||||
3537 CGCTATGCTACTATCTCTCAGTACCATATTATGCTCGATGGAGTGTGATGTTGGT 3596

QY 161 ArgGlySerLeuLeuMetLeuProCysProPheIleGlyArgLeuAsnPheCysGln 180
Db ::::|||||
3597 CGAGGCTCCCTGCTTATGCTCCCATGCTCCCTTATTTGGCGTTTGGAGCTTCTGCCAG 3656

QY 181 SerHisValIleLeuHisThrTyrCysGluHisMetAlaValValLysLeuAlaCysGly 200
Db ::::|||||
3657 AGCCACGTGATCTGTCACACATACGTGAGCAGATGCGCGTTGTGAAGCTGGCTGTGGA 3716

QY 201 AspThrArgProAsnArgValTyrGlyLeuThrAlaAlaLeuLeuValIleGlyValAsp 220
Db ::::|||||
3717 GACACACAGCCCTATCTGTTGTCAGGGCTCACAGCTGACATGTTGGTTCATCGGGTGTGAT 3776

QY 221 LeuPheCysIleGlyLeuSerTyrAlaLeuIleAlaGlnAlaValLeuArgLeuSerSer 240
Db ::::|||||
3777 TTGTTCTGCAATAGTCTCTCTATGCTCTCATGTGTACAGCTGCTTTCGTTCTCATCC 3836

QY 241 HisGluAlaArgSerLysAlaLeuGlyThrCysGlySerHisValCysValIleLeuIle 260
Db ::::|||||
3837 CATGAAGCTGTTCCAGGCCCTAGGACCTGTGTTCCCAATGCTGTGTTATCTCTCATC 3896

QY 261 SerTyrThrProAlaLeuPheSerPhePheThrHisArgPheGlyHisValProVal 280
Db ::::|||||
3897 TCTTACACACAGCTCTCTTTTCTTTTACATGCTTTGGCCACCATGTTCCACTC 3956

QY 281 HisIleHisIleLeuLeuAlaAsnValTyrLeuLeuLeuProProAlaLeuAsnProVal 300
Db ::::|||||
3957 CATATTACATTTCTTTGGCAATGTTTATCTGCTCTTCCACCTGCTCTTAACCTGTA 4016

QY 301 ValTyrGlyValLysThrLysGlnIleArgLysArgValValArgValPheGlnSerGly 320
Db ::::|||||
4017 GTATATGGGTTAAGACCAAGAGATTCGTGAAGGGTGTGTGAGAGTGTTCGAAGAGGA 4076

QY 321 GlnGlyMetGlyIleLysAlaSerGlu 329
Db ::::|||||
4077 CAGGGAATGGAATCAAGGCATCTGAG 4103

RESULT 15
AC146635/c
LOCUS
DEFINITION
Otolemur garnettii clone CH256-105A13, WORKING DRAFT SEQUENCE, 4
AC146635
AC146635.2 GI:37700288
VERSION
KEYWORDS
HTG; HTGS_PHASE2; HTGS_DRAFT.

```

SOURCE
ORGANISMREFERENCE
AUTHORS

Otolemur garnettii (small-eared galago)
Otolemur garnettii
Mammalia; Eutheria; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Strepsirhini; Galagonidae; Otolemur.
1 (bases 1 to 239335)
Antonellis, A., Ayele, K., Benjamin, B., Blakesley, R. W.,
Bouffard, G. G., Brinkley, C., Brooks, S., Chu, G., Coleman, B.,
Coleman, H., Engle, J., Granite, S., Guan, X., Gupta, J., Haghighi, P.,
Han, J., Hansen, N., Ho, S.-L., Hu, P., Hurle, B., Idol, J. R.,
Karls, E., Kwong, P., Laric, P., Larson, S., Lee-Lin, S.-O.,
Legaspi, R., Maduro, O. L., Maduro, V. B., Margulies, E. H., Masiello, C.,
Maskeri, B., McDowell, J., Mullikin, J. C., Paguirigan, C., Pearson, R.,
Portnoy, M. E., Prasad, A., Reddix-Dugue, N., Schandler, K.,
Schueler, M. G., Shah, K., Sison, C., Stantrop, S., Thomas, J. W.,
Thomas, P. J., Tsipouri, V., Vogt, J. L., Wetherby, K. D., Young, A. and
Green, E. D.

NISC Comparative Sequencing Initiative

Unpublished
2 (bases 1 to 239335)

Green, E. D.

Direct Submission

Submitted (11-SEP-2003) NIH Intramural Sequencing Center, 8717

Grovermont Circle, Gaithersburg, MD 20877, USA

3 (bases 1 to 239335)

Green, E. D.

Direct Submission

Submitted (17-OCT-2003) NIH Intramural Sequencing Center, 8717

Grovermont Circle, Gaithersburg, MD 20877, USA

On Oct 17, 2003 this sequence version replaced gi:34576313.

----- Genome Center

Center: NIH Intramural Sequencing Center

Center code: NISC

Web site: <http://www.nisc.nih.gov>

Contact: nisc.zoo@hgri.nih.gov

----- Project Information

Center project name: fgs

Center clone name: 105A13

The sequence data in this record represents an 'enhanced' version of a Phase 2 submission. Specifically, the indicated order and orientation of each sequence contig has been established using one or more of the following: read-pair data from individual subclones, overlaps with neighboring clones, alignment with available reference sequence (e.g., human), and/or confirmation by PCR testing. In addition, the sequence assembly is based on at least 8X average coverage in Q20 bases and has been reviewed to rule out gross misassemblies, the low-quality ends of sequence contigs have been trimmed away, and each base is associated with a Phrap-derived quality score.

----- Summary Statistics

Sequencing vector: plasmid; n/a; 100% of reads

Chemistry: Dye-terminator Big Dye; 100% of reads

Assembly program: Phrap; version 0.990319

Consensus quality: 238379 bases at least Q40

Consensus quality: 238694 bases at least Q30

Consensus quality: 238929 bases at least Q20

Insert size: 239000; agarose-fp

Insert size: 239035; sum-of-contigs

Quality coverage: 10.78x in Q20 bases; agarose-fp

Quality coverage: 10.78x in Q20 bases; sum-of-contigs

* NOTE: This is a 'working draft' sequence. It currently consists of 4 contigs. Gaps between the contigs are represented as runs of N. The order of the pieces is believed to be correct as given, however the sizes of the gaps between them are based on estimates that have been provided by the submitter.
* This sequence will be replaced
* by the finished sequence as soon as it is available and
* the accession number will be preserved.
* 1 4357: contig of 4357 bp in length
* 4358 4457: gap of unknown length

*	4458	66663: contig of 62206 bp in length
*	66664	66763: gap of unknown length
*	66764	232439: contig of 165676 bp in length
*	232440	232539: gap of unknown length
*	232540	239335: contig of 6796 bp in length.

FEATURES

```

source
1. 239335
/organism="Orculemur garnettii"
/mol_type="genomic DNA"
/db_xref="taxon:30611"
/clone="CH256-105A13"
/clone_lib="CH256"
1. 4357
/misc_feature
/notes="assembly_fragment
clone end:T7
vector_side:left"
4458. 66663
/notes="assembly_fragment"
66764. 232439
/notes="assembly_fragment"
232540. 239335
/notes="assembly_fragment
clone end:SP6
vector_side:right"

```

ORIGIN

Alignment Scores:			
Pred. No.:	3 59e-136	Length:	239335
Score:	158.00	Matches:	297
Percent Similarity:	90.22%	Conservative:	13
Best Local Similarity:	90.22%	Mismatches:	19
Query Match:	90.69%	Indels:	0
DB:	2	Gaps:	0

US-10-081-775-2 (1-329) X AC146635 (1-239335)

Qy	1	MetSerSerThrLeuGlyHisAsnMetGluSerProHisHisThrAspValAspProSer	20
Db	9997	ATATCCGGTACTCTTGGCCAGATCATGAATCTCTTAATACACACACATGGACCCCTCT	9938
Qy	21	ValPhePheLeuLeuGlyIleProGlyLeuGluGlnPheHisLeuTrpLeuSerLeuPro	40
Db	9937	ATTTCCTTCCTCTGGGTATCTCAGGCTAGAAAATTTCACTCTGTGGCTCTCACTTCCT	9878
Qy	41	ValCysGlyLeuGlyThrAlaThrIleValSglyAsnIleThrIleLeuValValValAla	60
Db	9877	GTGTGCTGCCTGGGCACAGCCACAGTGTGGGCAACATAACCATCTCTGGTTGTGTGGCC	9818
Qy	61	ThrGluProValLeuHisIleLysProValTyrLeuPheLeuCysMetLeuSerThrIleAsp	80
Db	9817	ACTGAGCCAGCCCTGCACAAAGCCTGTGTACTTTCTCTGTGCATGTTGTCAACCAATCGAC	9758
Qy	81	LeuAlaAlaSerValSerThrValProLysLeuLeuAlaIlePheTrpCysGlyAlaGly	100
Db	9757	TTGGCTCCCTCTGTCTCCACAGTTCCTCCCAAGCTACTGGGCATTCTCTGTGTGGAGCTGGA	9698
Qy	101	HisIleSerAlaSerAlaCysLeuAlaHisMetPhePheIleHisAlaPheCysMetMet	120
Db	9697	CATATATCTGCCTCTGTCTGGCTAGCACAGATGTTCTTTCATTATGCGCTTCTGCATGATG	9638
Qy	121	GluSerThrValLeuLeuAlaMetAlaPheAspArgTyrValAlaIleCysHisProLeu	140
Db	9637	GAGTCCACTGTGCTGTGGCCATGGCCTTTGATCGCTATGTGGCCATCTGCCATCCACTC	9578
Qy	141	ArgTyrAlaThrIleLeuThrAspThrIleIleAlaHisIleGlyValAlaAlaValVal	160
Db	9577	CGCTATGTCTACTATCTCTCACTGACACCAATTATTGCTCGCATTTGGAGTGGTAGCTATGGTG	9518
Qy	161	ArgGlySerLeuLeuMetLeuProCysProPhePheIleGlyArgLeuAsnPheCysGln	180
Db	9517	CGAGGCTCCCTGCTATGCTCCCCATGTCCTTCTTATTGGCGCTTTGAGCTCTGCCAG	9458
Qy	181	SerHisValIleLeuHisIleThrTyrCysGluHisMetAlaValValLysLeuAlaCysGly	200

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: August 27, 2004, 18:39:29 ; Search time 548 Seconds

(without alignments)
2954.652 Million cell updates/sec

Title: US-10-081-775-2

Perfect score: 1718

Sequence:

1 MSSTLGHNMESPHHTDVPDPS.....RKRVRVVFQSGQGMGIKASE 329

Scoring table:

BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 3237270 seqs, 2460713050 residues

Total number of hits satisfying chosen parameters: 6474540

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+p2n.model -DEV=xlh
-Q=/cgn2_1/USPTO.spool/US10081775/runat_20082004_170214_11184/app_query.fasta_1.519
-DB=PublishedApplications_NA -Q=FASTA -SUFFIX=rnpb -MINMATCH=0.1
-LOOFC=0 -LOOPEXT=0 -UNITS=bits -START=1 -END=-1 -MATRIX=blomsum62
-TRANS=human40.cdi -LIST=45 -DOCALIGN=200 -THR SCORE=pct -THR MAX=100
-MAXLEN=2000000000 -USER=US10081775 @CNC1.1.480 @runat_20082004_170214_11184
-NCPU=6 -ICPU=3 -NO MAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSPBLOCK=100
-LONGLOG -DRV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5
-XGAPOP=6 -XGAPEXT=7 -XGAPOP=10 -XGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : Published Applications NA.*

```
1: /cgn2_6/ptodata/2/pubpna/US07_PUBCOMB.seq.*
2: /cgn2_6/ptodata/2/pubpna/US07_PUBCOMB.seq.*
3: /cgn2_6/ptodata/2/pubpna/US06_NEW_PUB.seq.*
4: /cgn2_6/ptodata/2/pubpna/US06_PUBCOMB.seq.*
5: /cgn2_6/ptodata/2/pubpna/US07_NEW_PUB.seq.*
6: /cgn2_6/ptodata/2/pubpna/PCTUS_PUBCOMB.seq.*
7: /cgn2_6/ptodata/2/pubpna/US08_NEW_PUB.seq.*
8: /cgn2_6/ptodata/2/pubpna/US08_PUBCOMB.seq.*
9: /cgn2_6/ptodata/2/pubpna/US09A_PUBCOMB.seq.*
10: /cgn2_6/ptodata/2/pubpna/US09B_PUBCOMB.seq.*
11: /cgn2_6/ptodata/2/pubpna/US09C_PUBCOMB.seq.*
12: /cgn2_6/ptodata/2/pubpna/US09_NEW_PUB.seq.*
13: /cgn2_6/ptodata/2/pubpna/US09_NEW_PUB.seq.*
14: /cgn2_6/ptodata/2/pubpna/US10A_PUBCOMB.seq.*
15: /cgn2_6/ptodata/2/pubpna/US10B_PUBCOMB.seq.*
16: /cgn2_6/ptodata/2/pubpna/US10C_PUBCOMB.seq.*
17: /cgn2_6/ptodata/2/pubpna/US10_NEW_PUB.seq.*
18: /cgn2_6/ptodata/2/pubpna/US60_NEW_PUB.seq.*
19: /cgn2_6/ptodata/2/pubpna/US60_PUBCOMB.seq.*
```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	ID	Description
------------	-------	-------	--------	----	-------------

1	1718	100.0	1567	15	US-10-081-775-1	Sequence 1, Appli
2	1691	98.4	990	9	US-09-886-055-262	Sequence 262, App
3	1691	98.4	990	10	US-09-804-291-262	Sequence 262, App
4	1691	98.4	990	13	US-10-343-650A-353	Sequence 353, App
5	1691	98.4	998	15	US-10-025-806-35	Sequence 35, Appl
6	1691	98.4	998	15	US-10-025-806-37	Sequence 37, Appl
7	1691	98.4	1366	15	US-10-017-161-923	Sequence 923, App
8	1691	98.4	1390	16	US-10-292-798-797	Sequence 797, App
9	1649	96.0	966	17	US-10-297-021-42	Sequence 42, Appl
10	978.5	97.0	1300	16	US-10-292-798-833	Sequence 833, App
11	951.5	55.4	1242	15	US-10-017-161-959	Sequence 959, App
12	931.5	54.2	990	13	US-10-343-650A-315	Sequence 315, App
13	931.5	54.2	990	16	US-10-387-629-157	Sequence 157, App
14	931.5	54.2	1114	17	US-10-467-252-93	Sequence 93, Appl
15	931.5	54.2	1390	16	US-10-017-161-945	Sequence 945, App
16	931.5	54.2	1390	16	US-10-292-798-819	Sequence 819, App
17	920	53.6	1076	13	US-10-182-822A-23	Sequence 23, Appl
18	907	52.8	945	15	US-10-044-643-19	Sequence 19, Appl
19	904	52.6	936	9	US-09-886-055-114	Sequence 114, App
20	904	52.6	936	10	US-09-804-291-114	Sequence 114, App
21	904	52.6	936	13	US-10-343-650A-213	Sequence 213, App
22	904	52.6	985	15	US-10-044-643-15	Sequence 15, Appl
23	904	52.6	1001	15	US-10-220-382-31	Sequence 31, Appl
24	904	52.6	1258	15	US-10-017-161-909	Sequence 909, App
25	904	52.6	1336	16	US-10-292-798-783	Sequence 783, App
26	903	52.6	947	15	US-10-044-643-17	Sequence 17, Appl
27	901.5	52.5	1013	15	US-10-025-806-25	Sequence 25, Appl
28	899.5	52.4	1360	15	US-10-017-161-947	Sequence 947, App
29	899.5	52.4	1360	16	US-10-292-798-821	Sequence 821, App
30	889.5	51.8	957	9	US-09-886-055-102	Sequence 102, App
31	889.5	51.8	957	10	US-09-804-291-102	Sequence 102, App
32	889.5	51.8	957	13	US-10-343-650A-197	Sequence 197, App
33	889.5	51.8	990	15	US-10-025-806-1	Sequence 1, Appl
34	889.5	51.8	1345	15	US-10-017-161-877	Sequence 877, App
35	888.5	51.7	980	15	US-10-025-806-3	Sequence 3, Appl
36	886.5	51.6	982	11	US-09-844-861A-21	Sequence 21, Appl
37	886.5	51.6	1592	17	US-10-466-720-29	Sequence 29, Appl
38	885.5	51.5	966	9	US-09-886-055-260	Sequence 260, App
39	885.5	51.5	966	13	US-09-804-291-260	Sequence 189, App
40	885.5	51.5	966	13	US-10-343-650A-189	Sequence 189, App
41	885.5	51.5	971	16	US-10-024-212-91	Sequence 91, Appl
42	885.5	51.5	1246	15	US-10-017-161-925	Sequence 925, App
43	885.5	51.5	1366	16	US-10-292-798-799	Sequence 799, App
44	884.5	51.5	945	13	US-10-343-650A-159	Sequence 159, App
45	884.5	51.5	1151	11	US-09-844-861A-19	Sequence 19, Appl

ALIGNMENTS

RESULT 1
US-10-081-775-1

; Sequence 1, Application US/10081775

; Publication No. US2003060409A1

; GENERAL INFORMATION:

; APPLICANT: Bristol-Myers Squibb Company

; TITLE OF INVENTION: POLYNUCLEOTIDE ENCODING A NOVEL HUMAN G-PROTEIN COUPLED

; FILE REFERENCE: D0126 NP

; CURRENT APPLICATION NUMBER: US/10/081,775

; PRIOR FILING DATE: 2002-02-21

; PRIOR FILING DATE: 2001-02-21

; PRIOR APPLICATION NUMBER: US 60/270,134

; PRIOR FILING DATE: 2001-03-27

; NUMBER OF SEQ ID NOS: 69

; SOFTWARE: PatentIn version 3.0

; SEQ ID NO 1

; LENGTH: 1567

; TYPE: DNA

; ORGANISM: homo sapiens

; FEATURE:

; NAME/KEY: CDS

; LOCATION: (537) .. (1523)			
US-10-081-775-1			
Alignment Scores:			
Pred. No.:	1,42e-195	Length:	1567
Score:	1718.00	Matches:	329
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	100.00%	Indels:	0
DB:	15	Gaps:	0
US-10-081-775-2 (1-329) x US-10-081-775-1 (1-1567)			
Qy	1	MetSerSerThrLeuGlyHisAsnMetGluSerProHisHisThrAspValaspProSer	20
Db	537	ATGTCAGCACTCTTGGCCACAAACATGGAATCTCTCATCATCACTGATGTTGACCCCTTCT	596
Qy	21	ValPhePheLeuLeuGlyLeuProGlyLeuGluGlnPheHisLeuTrpLeuSerLeuPro	40
Db	597	GTCTTCTTCTCTGGGATCCAGGCTGGAACAAATTCATTTGTGGCTCTCACTCCCT	656
Qy	41	ValCysGlyLeuGlyThrAlaThrileValGlyAsnilleThrileLeuValValAla	60
Db	657	GTGTGTGCTTAGGCACGCCAATTTGGCAATATAACTATTCTGTGTTGTTGCC	716
Qy	61	ThrGluProValLeuHisLysProValTyrLeuPheLeuCysMetLeuSerThrileasp	80
Db	717	ACTGAACAGTCTTGACAAAGCCTGTGTACCTTTTCTGTGCATGCTCTCAACCATCGAC	776
Qy	81	LeuAlaAlaSerValSerThrValProLysLeuLeuAlaIlePheTrpCysGlyAlaGly	100
Db	777	TGGCTGCTCTGTCTCCACAGTCTCCCAAGCTACTGGCTATCTTCTGGTGGAGCCGA	836
Qy	101	HisIleSerAlaSerAlaCysLeuAlaHisMetPhePheIleHisAlaPheCysMetMet	120
Db	837	CATATATCTGCTCTGCTGCTGGCACATATGTTCTTCATTCATGCTCTGTCATGATG	896
Qy	121	GluSerThrValLeuLeuAlaMetAlaPheAspArgTyrValAlaIleCysHisProLeu	140
Db	897	GAGTCCACTGTGCTACTGCGCATGGCTTTGTATGCTAGCTAGTGGCCATCTGCCACCATC	956
Qy	141	ArgTyrAlaThrileLeuThrAspThrileAlaHisIleGlyValAlaAlaValVal	160
Db	957	CGCTATGCCAATCTCTACTGACACCATCATTGCCACATAGGGGTGGCAGCTGTAGTG	1016
Qy	161	ArgGlySerLeuLeuMetLeuProCysProPheIleGlyArgLeuAsnPheCysGln	180
Db	1017	CGAGGCTCCTCTCATGCTCCCATGCTCTTCTTATTGGGCGTTTGAATCTTGCCAA	1076
Qy	181	SerHisValIleLeuHisThrTyrCysGluHisMetAlaValValIleLeuAlaCysGly	200
Db	1077	AGCCATGTGATCCTACACAGTACTGTGAGCACTGGCTGTGGTGAAGCTGGCCTGTGA	1136
Qy	201	AspThrArgProAsnArgValTyrGlyLeuThrAlaAlaLeuLeuValIleGlyValasp	220
Db	1137	GACACAGGCCCTAACCGTGTATGGGCTGACAGCTGCACTGTGTGCTATGGGGTTGAC	1196
Qy	221	LeuPheCysIleGlyLeuSerTyrAlaLeuIleAlaGlnAlaValLeuArgLeuSerSer	240
Db	1197	TTGTTTGTGATGCTCTCTCTATGCTTAATGGCAAGTGTCTCTGCTCTCATCC	1256
Qy	241	HisGluAlaArgSerLysAlaLeuGlyThrCysGlySerHisValCysValIleLeuIle	260
Db	1257	CATGAAGCTCGGTCGAAGGCCCTAGGACCTGGTGTCCATGCTGTGTGTCATCCTCATC	1316
Qy	261	SerTyrThrProAlaLeuPheSerPhePheThrHisArgPheGlyHisHisValProVal	280
Db	1317	TCTTATACACAGCCCTCTCTCTCTTTTACACACCGCTTTGGCCATCAGCTTCAGTC	1376
Qy	281	HisIleHisIleLeuLeuAlaSerValTyrLeuLeuLeuProProAlaLeuAsnProVal	300
Db	1377	CATATTACATCTTTTGGCAATGTTATCTGCTTTTGGCACTGCTCTTATCTCTG	1436

Qy	301	ValTyrGlyValLysThrLysGlnIleArgLysArgValValArgValPheGlnSerGly	320
Db	1437	GTATATGGAGTTAAGACCAACACAGATCGTAAAGAGTGTTCAGGGTGTTCAGAGTGGG	1496
Qy	321	GlnGlyMetGlyIleLysAlaSerGlu	329
Db	1497	CAGGGAATGGCATCAAGGCATCTGAG	1523

RESULT 2

US-09-886-055-262

; Sequence 262, Application US/09886055

; Patent No. US20020132273A1

; GENERAL INFORMATION:

; APPLICANT: STRYER LUBERT

; APPLICANT: ZOZULYA, SERGEY

; TITLE OF INVENTION: RECEPTOR FINGERPRINTING, SENSORY PERCEPTION, AND

; TITLE OF INVENTION: BIOSENSORS OF CHEMICAL SENSANTS

; FILE REFERENCE: 078003-0277150

; CURRENT APPLICATION NUMBER: US/09/886,055

; CURRENT FILING DATE: 2001-06-22

; PRIOR APPLICATION NUMBER: 60/213,812

; PRIOR FILING DATE: 2000-06-22

; NUMBER OF SEQ ID NOS: 522

; SOFTWARE: Patent In Ver. 2.1

; SEQ ID NO 262

; LENGTH: 990

; TYPE: DNA

; ORGANISM: Homo sapiens

US-09-886-055-262

Alignment Scores:

Pred. No.:

Score:

Percent Similarity:

Best Local Similarity:

Query Match:

DB:

1,21e-192

1691.00

99.09%

98.78%

98.43%

9

Length:

Matches:

Conservative:

Mismatches:

Indels:

Gaps:

990

325

1

3

0

0

US-10-081-775-2 (1-329) x US-09-886-055-262 (1-990)

Qy	1	MetSerSerThrLeuGlyHisAsnMetGluSerProHisHisThrAspValaspProSer	20
Db	1	ATGTCAGCACTCTTGGCCACAAACATGGAATCTCTCAATACACATGATGTTGACCCCTCT	60
Qy	21	ValPhePheLeuLeuGlyLeuProGlyLeuGluGlnPheHisLeuTrpLeuSerLeuPro	40
Db	61	GTCTTCTTCTCTGGGATCCAGGCTGGAACAAATTCATTTGTGGCTCTCACTCCCT	120
Qy	41	ValCysGlyLeuGlyThrAlaThrileValGlyAsnilleThrileLeuValValAla	60
Db	121	GTGTGTGCTTAGGCACGCCAATTTGGGCAATATAACTATTCTGTGTTGTTGCC	180
Qy	61	ThrGluProValLeuHisLysProValTyrLeuPheLeuCysMetLeuSerThrileasp	80
Db	181	ACTGAACAGTCTTGACAAAGCCTGTGTACCTTTTCTGTGCATGCTCTCAACCATCGAC	240
Qy	81	LeuAlaAlaSerValSerThrValProLysLeuLeuAlaIlePheTrpCysGlyAlaGly	100
Db	241	TTGGCTGCTCTGTCTCCACAGTTCACAGCTACTGGCTATCTTCTGGTGTGAGCCGA	300
Qy	101	HisIleSerAlaSerAlaCysLeuAlaHisMetPhePheIleHisAlaPheCysMetMet	120
Db	301	CATATATCTGCTCTGCTGCTGGCACAGATGTTCTTCATTCATGCTCTGTCATGATG	360
Qy	121	GluSerThrValLeuLeuAlaMetAlaPheAspArgTyrValAlaIleCysHisProLeu	140
Db	361	GAGTCCACTGTGCTACTGCGCATGGCTTTGATCGCTAGCTGCTCTGCTCTGTCATGATG	420
Qy	141	ArgTyrAlaThrileLeuThrAspThrileAlaHisIleGlyValAlaAlaValVal	160
Db	421	CGCTATGCCAATCTCTCACTGACACCATCATTTGCCACATAGGGGTGGCAGCTGTAGTG	480
Qy	161	ArgGlySerLeuLeuMetLeuProCysProPhePheIleGlyArgLeuAsnPheCysGln	180

Db	481	CGAGGCTCCCTGCTCAATGCTCCCATGCTCCCTTTATTGGCGGTTGAACCTTCGCCAA	540
Qy	181	SerHisValIleLeuHisThrTyrCysGluHisMetAlaValValLysLeuAlaCysGly	200
Db	541	ASCCATGATCTACACACGTACTGTGAGCACATGGCTGTGGTGAAGCTGGCCTGTGA	600
Qy	201	AspThrArgProAsnArgValTyrGlyLeuThrAlaAlaLeuLeuValIleGlyValAsp	220
Db	601	GACACGAGCCCTAAACCGTGTGTATGGCTGACAGCTGCACGTGTGGTCAATGGGGTTGAC	660
Qy	221	LeuPheCysIleGlyLeuSerTyrAlaLeuIleAlaGlnAlaValLeuArgLeuSerSer	240
Db	661	TGTTTGGATTGGTCTCTCCATCGCTTAAGTGACACAGCTGTCTCGCTCTCATCC	720
Qy	241	HisGluAlaArgSerIleAlaLeuGlyThrCysGlySerHisValCysValIleLeuIle	260
Db	721	CATGAAGCTCGTCCAGGCCCTAGGACCTGTGGTTCCCATGCTGTCACTCATCT	780
Qy	261	SerTyrThrProAlaLeuPheSerPhePheThrHisArgPheGlyHisHisValProVal	280
Db	781	TCTTATACACGAGCCCTCTCTCTCTTTTACACACCCGCTTTGGCCATCACGTTCCAGTC	840
Qy	281	HisIleHisIleLeuLeuAlaAsnValTyrLeuLeuLeuProProAlaLeuAsnProVal	300
Db	841	CATATTCACTTCTTTGGCCAAATGTTTATCTGCTTTTGGCCACCTGCTCTTAATCCTGTG	900
Qy	301	ValTyrGlyValLysThrLysGlnIleArgLysArgValValArgValPheGlnSerGly	320
Db	901	GTATATGAGATTAGACCAACACAGATCCGTAACAGAGTTGTCAAGGTGTTTCAAAGTGGG	960
Qy	321	GlnGlyMetGlyIleLysAlaSerGlu	329
Db	961	CAGGGAATGGGCATCAAGGCATCTGAG	987

RESULT 3

```

US-09-804-291-262
; Sequence 262, Application US/09804291
; Publication No. US20030088059A1
; GENERAL INFORMATION:
; APPLICANT: ZOZULA, SERGEY
; TITLE OF INVENTION: HUMAN OLFACTORY RECEPTORS AND GENES ENCODING SAME
; FILE REFERENCE: P 0378005
; CURRENT APPLICATION NUMBER: US/09/804,291
; CURRENT FILING DATE: 2001-03-13
; PRIOR APPLICATION NUMBER: 60/188,914
; PRIOR FILING DATE: 2000-03-13
; PRIOR APPLICATION NUMBER: 60/192,033
; PRIOR FILING DATE: 2000-03-24
; PRIOR APPLICATION NUMBER: 60/198,474
; PRIOR FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/199,335
; PRIOR FILING DATE: 2000-04-24
; PRIOR APPLICATION NUMBER: 60/207,702
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/213,849
; PRIOR FILING DATE: 2000-06-23
; PRIOR APPLICATION NUMBER: 60/226,534
; PRIOR FILING DATE: 2000-08-16
; PRIOR APPLICATION NUMBER: 60/230,732
; PRIOR FILING DATE: 2000-09-07
; PRIOR APPLICATION NUMBER: 60/266,862
; PRIOR FILING DATE: 2001-02-07
; NUMBER OF SEQ ID NOS: 529
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 262
; LENGTH: 990
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-804-291-262

Alignment Scores:
Pred. No.: 1.121e-192 Length: 990

```

Db 961 CAGGGAATGGGCATCAAGGCATCTTGAG 987

RESULT 4

US-10-343-650A-353

Sequence 353, Application US/10343650A

Publication No. US20040067499A1

GENERAL INFORMATION:

APPLICANT: HAGA, TATSUYA

TITLE OF INVENTION: NOVEL G-PROTEIN COUPLED RECEPTOR

FILE REFERENCE: 31671-186347

CURRENT APPLICATION NUMBER: US/10/343,650A

CURRENT FILING DATE: 2003-07-21

PRIOR APPLICATION NUMBER: JP 2000/237818

PRIOR FILING DATE: 2000-08-04

PRIOR APPLICATION NUMBER: JP 2001/34434

PRIOR FILING DATE: 2001-02-13

NUMBER OF SEQ ID NOS: 694

SOFTWARE: PatentIn Ver. 2.1

SEQ ID NO 353

LENGTH: 990

TYPE: DNA

ORGANISM: Homo sapiens

FEATURE:

NAME/KEY: CDS

LOCATION: (1)..(990)

US-10-343-650A-353

Alignment Scores:

Pred. No.: 1,21e-192 Length: 990

Score: 1691.00 Matches: 325

Percent Similarity: 99.09% Conservative: 1

Best Local Similarity: 98.78% Mismatches: 3

Query Match: 98.43% Indels: 0

DB: 13 Gaps: 0

US-10-081-775-2 (1-329) x US-10-343-650A-353 (1-990)

Qy 1 MetSerThrLeuGlyHisAsnMetGluSerProHisHisThrAspValAspProSer 20

Db 1 ATGTCCAGCACTCTGTGGCCACACATGGAAATCTTCCTAAATCACACATGATGTGACCTTCT 60

Qy 21 ValPhePheLeuGlyIleProGlyLeuGluGlnPheHisLeuTrpLeuSerLeuPro 40

Db 61 GTCTTCTCTCTGCTGGGCATCCAGGCTCTGGAAACAATTTCAATTTGGTCTCTCACTCCCT 120

Qy 41 ValCysGlyLeuGlyThrAlaThrIleValGlyAsnIleThrIleLeuValValAla 60

Db 121 GTGTGTGGCTTAGGCACAGCCAAATGTGGGCAATATAACTATTCTGTGTGTGTGGTGC 180

Qy 61 ThrGluProValLeuHisLysProValTyrrPheLeuCysMetLeuSerThrIleAsp 80

Db 181 ACTGAACCACTGTGCACAGCCCTGTGTACCTTTCTGTGCATGTCTCTCAACCATCGAC 240

Qy 81 LeuAlaIleSerValSerThrValProLysLeuLeuAlaIlePheTrpCysGlyAlaGly 100

Db 241 TTGGCTGCCTCTGTCTCCACAGTITCCCAAGCTACTGGCTATCTCTCTGTGGTGGAGCCGA 300

Qy 101 HistSerAlaSerAlaCysLeuAlaHisMetPheIleHisAlaPheCysMetMet 120

Db 301 CATATATCTGCCTCTGTCTCCACAGTITCCCAAGCTACTGGCTATCTCTCTGTGGTGGAGCCGA 360

Qy 121 GluSerThrValLeuLeuAlaMetAlaPheAspArgTyrrValAlaIleCysHisProLeu 140

Db 361 GAGTCCACTGTGTACTGGCCATGGCTTTGATCGCTACGTGGCCATCTGCCACCCACTC 420

Qy 141 ArgTyrrAlaThrIleLeuThrAspThrIleIleAlaHisIleGlyValAlaAlaVal 160

Db 421 CGTATGTCACAAATCTCTCACTGACACCATCATTTCCCAATAGGGTGGAGCTGTAGTG 480

Qy 161 ArgGlySerLeuLeuMetLeuProCysProPheIleGlyArgLeuAsnPheCysGln 180

Db 481 CTAGGCTCTCCGCTATCTATCTCCATCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 540

; PRIOR APPLICATION NUMBER: 60/261,498
 ; PRIOR FILING DATE: 2001-01-12
 ; PRIOR APPLICATION NUMBER: 60/263,689
 ; PRIOR FILING DATE: 2001-01-24
 ; PRIOR APPLICATION NUMBER: 60/276,464
 ; PRIOR FILING DATE: 2001-02-08
 ; PRIOR APPLICATION NUMBER: 60/271,021
 ; PRIOR FILING DATE: 2001-02-22
 ; PRIOR APPLICATION NUMBER: 60/275,946
 ; PRIOR FILING DATE: 2001-03-14
 ; PRIOR APPLICATION NUMBER: 60/278,150
 ; PRIOR FILING DATE: 2001-03-23
 ; PRIOR APPLICATION NUMBER: 60/285,718
 ; PRIOR FILING DATE: 2001-04-23
 ; PRIOR APPLICATION NUMBER: 60/312,902
 ; PRIOR FILING DATE: 2001-08-16
 ; PRIOR APPLICATION NUMBER: 60/257,876
 ; PRIOR FILING DATE: 2000-12-21
 ; PRIOR APPLICATION NUMBER: 60/260,718
 ; PRIOR FILING DATE: 2001-01-10
 ; PRIOR APPLICATION NUMBER: 60/284,591
 ; PRIOR FILING DATE: 2001-04-18
 ; NUMBER OF SEQ ID NOS: 352
 ; SOFTWARE: PatentIn Ver. 2.1
 ; SEQ ID NO 35
 ; LENGTH: 998
 ; TYPE: DNA
 ; ORGANISM: Homo sapiens
 ; FEATURE:
 ; NAME/KEY: CDS
 ; LOCATION: (2)..(998)
 ; US-10-025-806-35

Alignment Scores:

Pred. No.: 1,22e-192 Length: 998
 Score: 1691.00 Matches: 325
 Percent Similarity: 99.09% Conservative: 1
 Best Local Similarity: 98.78% Mismatches: 3
 Query Match: 98.43% Indels: 0
 DB: 15 Gaps: 0

US-10-081-775-2 (1-329) x US-10-025-806-35 (1-998)

Qy 1 MetSerThrLeuGlyHisAsnMetGluSerProHisHisThrAspValAspProSer 20
 Db 2 ATGTCAGACCTCTTGGCCACAAATGGAATCTCTTAATCACATGATGTTGACCTTCT 61
 Qy 21 ValPhePheLeuLeuGlyIleProGlyLeuGluGlnPheHisLeuTrpLeuSerLeuPro 40
 Db 62 GTCTTCTTCTCTCTGGGCATCCAGGCTCTGGACAAATTTCAATTTGGCTCTCACTCCCT 121
 Qy 41 ValCysGlyLeuGlyThrAlaThrIleValGlyAsnIleThrIleLeuValValAla 60
 Db 122 GTGTGTGGCTTAGGCACAGCCACAAATTTGGGCAATATAACTATTTGGTTGTTGGCC 181
 Qy 61 ThrGluProValLeuHisLysProValTyrLeuPheLeuCysMetLeuSerThrIleAsp 80
 Db 182 ACTGAACAGCTGTGCACAGCCTGTGTACCTTTTCTGTGCATGCTCTCAACCATCGAC 241
 Qy 81 LeuAlaAlaSerValSerThrValProLysLeuLeuAlaIlePheTrpCysGlyAlaGly 100
 Db 242 TTGGCTGCCTCTGTCTCCACAGTTCACAGCTACTGGCTATCTTCTGGTGTGGAGCCGA 301
 Qy 101 HistSerAlaSerAlaCysLeuAlaHisMetPhePheIleHisAlaPheCysMetMet 120
 Db 302 CATATATCTGCCTCTGCCTGCCTGGCAGAGATGTTCTTCAATTCATGCTTCTTGCATGATG 361
 Qy 121 GluSerThrValLeuLeuAlaMetAlaPheAspArgTyrValAlaIleCysHisProLeu 140
 Db 362 GAGTCCACTGTGCTACTGGCCATGGCCCTTTGATCGCTACGTGGCCATCTGCCACCCACTC 421
 Qy 141 ArgTyrAlaThrIleLeuThrAspThrIleIleAlaHisIleGlyValAlaValVal 160

Db 422 CGCTATGCCACAATCTCTACTGACACCATCATATTGCCACATAGGGGTGGCAGCTGTAGTG 481
 Qy 161 ArgGlySerLeuLeuMetLeuProCysProPhePheIleGlyArgLeuAsnPheCysGln 180
 Db 482 CGAGGCTCCTCTCATCTCCCATGTCCCTTCTTATTGGGCGTTTGAACTTCTGCCAA 541
 Qy 181 SerHisValIleLeuHisThrTyrCysGluHisMetAlaValValLysLeuAlaCysGly 200
 Db 542 AGCCATGTGATCTACACAGCTACTGTGAGCACATGGCTGTGTGAAGCTGGCTGTGGA 601
 Qy 201 AspThrArgProAsnArgValTyrGlyLeuThrAlaAlaLeuValIleGlyValAsp 220
 Db 602 GACACAGCCCTAACCGTGTGTATGGCTGTGACAGCTGCACCTGTGTCAATGGGGTTGAC 661
 Qy 221 LeuPheCysIleGlyLeuSerTyrAlaLeuIleAlaGlnAlaValLeuArgLeuSerSer 240
 Db 662 TTGTTTTCATTTGGTCTCTCTCTATGCCCTAAGTCACAGCTGTCTTTCGCTCTCATCC 721
 Qy 241 HisGluAlaArgSerLysAlaLeuGlyThrCysGlySerHisValCysValIleLeuIle 260
 Db 722 CATGAAGCTCGGTCCAGGCCCTTAGGGACCTGTGGTTCCTCATGTCTGTGTATCCTCATC 781
 Qy 261 SerTyrThrProAlaLeuPheSerPheThrHisArgPheGlyHisHisValProVal 280
 Db 782 TCTTATACACAGCCCTCTTCTCTTTTACACACCCGCTTTGGCCATCACGTTCCAGTC 841
 Qy 281 HistLeuIleLeuLeuAlaAsnValTyrLeuLeuLeuProAlaLeuAsnProVal 300
 Db 842 CATATTACATTTCTTTGGCCAAATGTTTATCTGCTTTTGGCCACCTGCTCTTAATCTGTG 901
 Qy 301 ValTyrGlyValLysThrLysGlnIleArgLysArgValValArgValPheGlnSerGly 320
 Db 902 GTATATGGAGTTAAGACCAACACAGATCCGTAAGAGTGTTCAGGGTGTCTTCAAGTGGG 961
 Qy 321 GlnGlyMetGlyIleLysAlaSerGlu 329
 Db 962 CAGGGAATGGGCATCAGGCCATCTGAG 988

RESULT 6

US-10-025-806-37
 ; Sequence 37, Application US/10025806
 ; Publication No. US20030198955A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Li, Li
 ; APPLICANT: Padigaru, Muralidhara
 ; APPLICANT: Ballinger, Robert
 ; APPLICANT: Kekuda, Ramesh
 ; APPLICANT: Colman, Steven
 ; APPLICANT: Spytek, Kimberly
 ; APPLICANT: Casman, Stacie
 ; APPLICANT: Edinger, Shlomit
 ; APPLICANT: Gerlach, Valerie
 ; APPLICANT: Sciore, Paul
 ; APPLICANT: Smithson, Glenda
 ; APPLICANT: Peyman, John
 ; APPLICANT: MacDougall, John
 ; APPLICANT: Stone, David
 ; APPLICANT: Vernet, Corine
 ; APPLICANT: Shenoy, Suresh
 ; APPLICANT: Gunther, Erik
 ; APPLICANT: Miller, Isabelle
 ; APPLICANT: Tchernev, Velizar
 ; APPLICANT: Anderson, David
 ; APPLICANT: Gusev, Vladimir
 ; APPLICANT: Malyankar, Uriel
 ; APPLICANT: Zhong, Haihong
 ; APPLICANT: Ellerman, Karen
 ; APPLICANT: Wolenc, Adam
 ; TITLE OF INVENTION: NOVEL PROTEINS AND NUCLEIC ACIDS ENCODING SAME
 ; FILE REFERENCE: 21402-224 AB
 ; CURRENT APPLICATION NUMBER: US/10/025,806
 ; CURRENT FILING DATE: 2001-12-19
 ; PRIOR APPLICATION NUMBER: 60/256,635

Pred. No.: 1.99e-192 Length: 1366
Score: 1691.00 Matches: 325
Percent Similarity: 99.09% Conservatives: 1
Best Local Similarity: 98.78% Mismatches: 3
Query Match: 98.43% Indels: 0
DB: 15 Gaps: 0

US-10-081-775-2 (1-329) x US-10-017-161-923 (1-1366)

QY 1 MetSerSerThrLeuGlyHisAsnMetGluSerProHisHisThrAspValAspProSer 20
DB 177 ATGTCAGACACTCTTGGCCACACATGGAATCTCTTAATCACACTGATGTGACCTTCT 236
QY 21 ValPhePheLeuLeuGlyLeuProGlyLeuGluGlnPheHisLeuTrpLeuSerLeuPro 40
DB 237 GTCCTTCTCTCTGGGCATCCAGGTCTGGACAAATTCATTTGGCTCTCACTCCCT 296
QY 41 ValCysGlyLeuGlyThrAlaThrIleValGlyAsnIleThrIleLeuValValAla 60
DB 297 GTGTGTGGCTTAGGCACAGCCCAATTTGGGCAATATAACTATCTGTTGTTGGC 356
QY 61 ThrGluProValLeuHisLysProValTyrLeuPheLeuCysMetLeuSerThrIleAsp 80
DB 357 ACTGAACACAGCTTTGCACACAGCTGTGTACCTTTTCTGTGATGCTCTCAACCATCGAC 416
QY 81 LeuAlaAlaSerValSerThrValProLysLeuLeuAlaIlePheTrpCysGlyAlaGly 100
DB 417 TTGGTGTCTGTCTCCACAGTTCACAGCTACTGGCTATCTTCTGGTGTGGAGCCGGA 476
QY 101 HisIleSerAlaSerAlaCysLeuAlaHisMetPhePheIleHisAlaPheCysMetMet 120
DB 477 CATATATCTGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 536
QY 121 GluSerThrValLeuLeuAlaMetAlaPheAspArgTyrValAlaIleCysHisProLeu 140
DB 537 GAGTCCACTGTGTACTGGCCATGGCTTTGATCGCTAGTGGCCATCTGCGCCACCTC 596
QY 141 ArgTyrAlaThrIleLeuThrAspThrIleIleAlaHisIleGlyValAlaValVal 160
DB 597 CGCTATGACCAATCTCTACAGACACATCTGAGCAGCATGGCTGTGTAAGCTGGCTG 656
QY 161 ArgGlySerLeuLeuMetLeuProCysProPhePheIleGlyArgLeuAsnPheCysGln 180
DB 657 CGAGCTCCCTGCTCATGCTCCCATGCTCCCTTCTTATTGGGCGTTGAATCTTGCCAA 716
QY 181 SerHisValIleLeuHisThrTyrCysGluHisMetAlaValValValLysLeuAlaCysGly 200
DB 717 AGCCATGTGATCTTACACACACTACTGTGAGCAGCATGGCTGTGTAAGCTGGCTG 776
QY 201 AspThrArgProAsnArgValTyrGlyLeuThrAlaAlaLeuValIleGlyValAsp 220
DB 777 GACACACAGGCTAACCGTGTGTATGGGCTGACAGCTGCTGTTGGTCAITGGGGTTGAC 836
QY 221 LeuPheCysIleGlyLeuSerTyrAlaLeuIleAlaGlnAlaValLeuArgLeuSerSer 240
DB 837 TTGTTTGTGCATTTGCTCTCTATGCTCCCTAAGTGCACAGCTGCTCTGCTCTCATCC 896
QY 241 HisGluAlaArgSerLysAlaLeuGlyThrCysGlySerHisValCysValIleLeuIle 260
DB 897 CATGAAGCTGGTCCAAAGGCTTAGGACCTGTGGTCCCATGCTCTGTGCTCATCTCATC 956
QY 261 SerTyrThrProAlaLeuPheSerPhePheThrHisArgPheGlyHisHisValProVal 280
DB 957 TCTTATACACAGCCCT 1016
QY 281 HisIleHisIleLeuLeuAlaAsnValTyrIleLeuLeuProProAlaLeuAsnProVal 300
DB 1017 CATATTACATCTTTTGGCAATGTTTATCTGCTTTTGGACCTGCTCTTAATCTCTGTG 1076
QY 301 ValTyrGlyValLysThrLysGlnIleArgLysArgValValArgValPheGlnSerGly 320
DB 1077 GTATATGGAGTTAAGACCAACAGATCCGTAAGAGTTGTGAGGGTGTTCAAAGTGGG 1136

QY 321 GlnGlyMetGlyIleLysAlaSerGlu 329
DB 1137 CAGGGAATGGCATCAAGGCATCTGAG 1163

RESULT 8

US-10-292-798-797
; Sequence 797, Application US/10292798
; Publication No. US2003023583A1
; GENERAL INFORMATION:
; APPLICANT: SUWA, MAKIKO
; APPLICANT: ASAI, KIYOSHI
; APPLICANT: AKIYAMA, YUTAKA
; APPLICANT: ABURATANI, HIROYUKI
; TITLE OF INVENTION: GUANOSINE TRIPHOSPHATE-BINDING PROTEIN COUPLED RECEPTORS
; FILE REFERENCE: 084335/166
; CURRENT APPLICATION NUMBER: US/10/292,798
; CURRENT FILING DATE: 2002-11-13
; PRIOR APPLICATION NUMBER: 10/017,161
; PRIOR FILING DATE: 2001-12-18
; PRIOR APPLICATION NUMBER: JP 2001-246789
; NUMBER OF SEQ ID NOS: 2070
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 797
; LENGTH: 1390
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; LOCATION: source
; FEATURE:
; LOCATION: (1)..(1390)
; NAME/KEY: CDS
; LOCATION: (201)..(1190)
US-10-292-798-797

Alignment Scores:
Pred. No.: 2.05e-192 Length: 1390
Score: 1691.00 Matches: 325
Percent Similarity: 99.09% Conservatives: 1
Best Local Similarity: 98.78% Mismatches: 3
Query Match: 98.43% Indels: 0
DB: 16 Gaps: 0

US-10-081-775-2 (1-329) x US-10-292-798-797 (1-1390)

QY 1 MetSerSerThrLeuGlyHisAsnMetGluSerProHisHisThrAspValAspProSer 20
DB 201 ATGTCAGACACTCTTGGCCACACATGGAATCTCTTAATCACACTGATGTGACCTTCT 260
QY 21 ValPhePheLeuLeuGlyIleProGlyLeuGluGlnPheHisLeuTrpLeuSerLeuPro 40
DB 261 GTCTTCTCTCTCTGGGCATCCAGGTCTGGACAAATTCATTTGGCTCTCACTCCCT 320
QY 41 ValCysGlyLeuGlyThrAlaThrIleValGlyAsnIleThrIleLeuValValAla 60
DB 321 GTGTGTGGCTTAGGCACAGCCCAATTTGGGCAATATAACTATCTTGTGGTGTGGCC 380
QY 61 ThrGluProValLeuHisLysProValTyrLeuPheLeuCysMetLeuSerThrIleAsp 80
DB 381 ACTGAACAGCTTTGGCAAGCTGTGTACCTTTTCTGTGATGCTCTCAACCATCGAC 440
QY 81 LeuAlaAlaSerValSerThrValProLysLeuLeuAlaIlePheTrpCysGlyAlaGly 100
DB 441 TTGGTGTCTCTGTCTCCACAGTTCACAGTCTCCAGCTACTGGCTATCTTCTGGTGTGGAGCCGGA 500
QY 101 HisIleSerAlaSerAlaCysLeuAlaHisMetPhePheIleHisAlaPheCysMetMet 120
DB 501 CATATATCTGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 560
QY 121 GluSerThrValLeuLeuAlaMetAlaPheAspArgTyrValAlaIleCysHisProLeu 140
DB 561 GAGTCCACTGTGTACTTGGCCATGGCTTTGATGCTAGCTGGCCATCTGCCACCCACTC 620

```
Qy 141 ArgTyrAlaThrIleLeuThrAspThrIleleAlaHisIleGlyValAlaAlaValVal 160
Db 621 CGCTATGCGCAAACTCTCACTGACACCATCATTTGCCACATAGGGGTGGCAGCTGTAGTG 680
Qy 161 ArgGlySerLeuLeuMetLeuProCysProPhePheIleGlyArgLeuAsnPheCysGln 180
Db 681 CGAGGCTCCCTCGCTCATGCTCCATGCTCCCTTCTTATGGGCGGTTTGAACCTTCGCCAA 740
Qy 181 SerHisValIleLeuHisThrTyrCysGluHisMetAlaValValValLeuAlaCysGly 200
Db 741 AGCCATGTGATCCACACAGTACTGTGACACATGGCTGTGGTGAAGCTGGCCTGTGGA 800
Qy 201 AspThrArgProAsnArgValTyrGlyLeuThrAlaAlaLeuLeuValIleGlyValAsp 220
Db 801 GACACGAGCGCTAACCGGTGTATGGGCTGACAGCTGCACCTGTGGTCAATTTGGGTTGAC 860
Qy 221 LeuPheCysIleGlyLeuSerTyrAlaLeuIleAlaGlnAlaValLeuArgLeuSerSer 240
Db 861 TTGTTTTGCAATGGTCTCTCTATGCCCTTAAGTGACAGCTGTCTTGGCCTCTCATCC 920
Qy 241 HisGluAlaArgSerLysAlaLeuGlyThrCysGlySerHisValCysValIleLeuIle 260
Db 921 CATGAAGCTCGTCCAGGCGCTAGGACCTGTGGTTCCTCATGTCTGTGTCATCCTCATC 980
Qy 261 SerTyrThrProAlaLeuPheSerPhePheThrHisArgPheGlyHisHisValProVal 280
Db 981 TCTTATACACACCGCTCTCTCTCTTTTATACACACCGCTTTGGCCATCAGCTTCCAGTC 1040
Qy 281 HisIleHisIleLeuLeuAlaAsnValTyrLeuLeuLeuProProAlaLeuAsnProVal 300
Db 1041 CATATTCATATCTTTTGGCCAAATGTTTATCTGCTTTTGGCCACCTGCTCTTAACTCTGTG 1100
Qy 301 ValTyrGlyValIysThrIysGlnIleArgIysArgValValArgValPheGlnSerGly 320
Db 1101 GTATATGAGTAAAGCAACACAGATCCGTAAGAGAGTTGTGAGGGTGTTCAGAGTGGG 1160
Qy 321 GlnGlyMetGlyIleLysAlaSerGlu 329
Db 1161 CAGGGAATGGCATCAAGGCATCTGAG 1187

RESULT 9
US-10-297-021-42
; Sequence 42, Application US/10297021
; Publication NO. US20040023294A1
; GENERAL INFORMATION:
; APPLICANT: INCYTE GENOMICS, INC.
; APPLICANT: PATTERSON, Chandra
; APPLICANT: TRIBOULEY, Catherine M.
; APPLICANT: YAO, Monique G.
; APPLICANT: GRIFFIN, Jennifer A.
; APPLICANT: THORNTON, Michael
; APPLICANT: LU, Yan
; APPLICANT: KALLICK, Deborah A.
; APPLICANT: GANDHI, Ameena R.
; APPLICANT: AU-YOUNG, Janice
; TITLE OF INVENTION: G-PROTEIN COUPLED RECEPTORS
; FILE REFERENCE: SF-0781 PCT
; CURRENT APPLICATION NUMBER: US/10/297,021
; CURRENT FILING DATE: 2001-05-22
; PRIOR APPLICATION NUMBER: 60/206,222; 60/207,476; 60/208,834; 60/208,861; 60/209,868
; PRIOR FILING DATE: 2000-05-22; 2000-05-25; 2000-06-02; 2000-06-02; 2000-06-07
; NUMBER OF SEQ ID NOS: 46
; SOFTWARE: PERL Program
; SEQ ID NO 42
; LENGTH: 966
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; OTHER INFORMATION: Incyte ID No. US20040023294A1 7476077CB1
US-10-297-021-42
```

Alignment Scores:

```
Pred. No.: 1,28e-187 Length: 966
Score: 1649.00 Matches: 317
Percent Similarity: 99.07% Conservative: 1
Best Local Similarity: 98.75% Mismatches: 3
Query Match: 95.98% Indels: 0
DB: 17 Gaps: 0

US-10-081-775-2 (1-329) x US-10-297-021-42 (1-966)

Qy 9 MetGluSerProHisHisThrAspValAspProSerValPhePheLeuLeuGlyIlePro 28
Db 1 ATGGAATCTCCTAATACACATGATGTTGACCCCTCTTGTCTCTCTCTCTCTCTCTCTCA 60
Qy 29 GlyLeuGluGlnPheHisLeuThrLeuSerLeuProValCysGlyLeuGlyThrAlaThr 48
Db 61 GGTCTGGAAACAATTTCAATTTGTGGCTCTCACTCCCTGTGTGGCTTAGGCACAGCCACA 120
Qy 49 IleValGlyAsnIleThrIleLeuValValAlaThrGluProValLeuHisLysPro 68
Db 121 ATGTGGGCAATATAACTATTCTGGTGTGTGGCACTGAACCACTCTTGCAAGGCT 180
Qy 69 ValTyrLeuPheLeuCysMetLeuSerThrIleAspLeuAlaAlaSerValSerThrVal 88
Db 181 GTGTACCTTTTCTGTGCATGCTCTCAACCATCGACTTGGCTGCTCTGTCTCCACAGTT 240
Qy 89 ProLysLeuLeuAlaIlePheThrCysGlyAlaGlyHisIleSerAlaSerAlaCysLeu 108
Db 241 CCCAAGCTACTGGCTATCTTCTGGTGTGGAGCCGACATATATCTGCTCTGCTCTGCTG 300
Qy 109 AlaHisMetPhePheIleHisAlaPheCysMetMetGluSerThrValLeuLeuAlaMet 128
Db 301 GCACAGATGTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 360
Qy 129 AlaPheAspArgTyrValAlaIleCysHisProLeuArgTyrAlaThrIleLeuThrAsp 148
Db 361 GCCTTTGATCGCTAGCTGGCCATCTGCCACCCACATCCGCTATGCCACAAATCTCTCACTG 420
Qy 149 ThrIleIleAlaHisIleGlyValAlaAlaValArgGlySerLeuLeuMetLeuPro 168
Db 421 ACCATCATGCCCAATAGGGGTGGAGCTGTAGTGGAGGGCTCTCTGCTCATGCTCCCA 480
Qy 169 CysProPhePheIleGlyArgLeuAsnPheCysGlnSerHisValIleLeuHisThrTyr 188
Db 481 TGTCTCTCTCTTATTGGGCGTTTGAACCTTCTGCCAAAGCCATGTGATCTCTACACAGTAC 540
Qy 189 CysGluHisMetAlaValIlyLeuAlaCysGlyAspThrArgProAsnArgValTyr 208
Db 541 TGTGAGCACATGGCTGTGGTGAAGCTGGCTGTGGAGACACAGGCGCTTAACCGTGTGTAT 600
Qy 209 GlyLeuThrAlaAlaLeuLeuValIleGlyValAspLeuPheCysIleGlyLeuSerTyr 228
Db 601 GGGCTGACAGCTGCATGTTGGTTCATCTGGGGTGTGACTTGTTCATTCGCTCTCTCTAT 660
Qy 229 AlaLeuIleAlaGlnAlaValLeuArgLeuSerSerHisGluAlaAlaArgSerLysAlaLeu 248
Db 661 GCCCTAAGTGCAACAGCTGCTCTTGGCTCTCATCCATGAAGCTCGTCCAAGGCCCTA 720
Qy 249 GlyThrCysGlySerHisValCysValIleLeuLeuSerTyrThrProAlaLeuPheSer 268
Db 721 GGGACCTGTGGTTCCTCATCTGTGTGCATCTCTTATACACAGCCCTCTCTCTCTCT 780
Qy 269 PhePheThrHisArgPheGlyHisHisValProValHisIleHisIleLeuLeuAlaAsn 288
Db 781 TTTTITACACACCGCTTGGCCATCAGCTTCCAGTCCATATTCACATCTCTTTTGGCCAAT 840
Qy 289 ValTyrLeuLeuLeuProProAlaLeuAsnProValValTyrGlyValIysThrLysGln 308
Db 841 GTTTATCTGCTTTTGGCACCTGCTCTTAATCTCTGTGTATATGGAGTTAAGACCAACAG 900
Qy 309 IleArgIysArgValValArgValPheGlnSerGlyGlnGlyMetGlyIleLysAlaSer 328
Db 901 ATCCGTAAGAGAGTTGTCTAGGGTGTTCAGAGTGGGCAAGGGAATGGGCATCAAGGCATCT 960
```


QY	329	Glu 329	Db	489	GAATCTGGTATACCTAGCAATGGCCTTTGACCGCTACTTAGCCATTGCTGGCCTCTG	548
Db	961	GAG 963	QY	141	ArgTyrAlaThrIleLeuThrAspThrIleIleAlaHisIleGlyValAlaValVal	160
RESULT 10						
US-10-292-798-833						
; Sequence 833, Application US/10292798						
; Publication No. US20030235833A1						
; GENERAL INFORMATION:						
; APPLICANT: SUWA, MAKIKO						
; APPLICANT: ASAI, KIYOSHI						
; APPLICANT: AKIYAMA, YUTAKA						
; APPLICANT: ABURATANI, HIROYUKI						
; TITLE OF INVENTION: GUANOSINE TRIPHOSPHATE-BINDING PROTEIN COUPLED RECEPTORS						
; FILE REFERENCE: 084335/166						
; CURRENT APPLICATION NUMBER: US/10/292,798						
; PRIOR FILING DATE: 2002-11-13						
; PRIOR APPLICATION NUMBER: 10/017,161						
; PRIOR FILING DATE: 2001-12-18						
; PRIOR APPLICATION NUMBER: JP 2001-246789						
; PRIOR FILING DATE: 2001-06-18						
; NUMBER OF SEQ ID NOS: 2070						
; SOFTWARE: PatentIn Ver. 2.1						
; SEQ ID NO 833						
; LENGTH: 1300						
; TYPE: DNA						
; ORGANISM: Homo sapiens						
; FEATURE: source						
; LOCATION: (1)..(1300)						
; NAME/KEY: CDS						
; LOCATION: (201)..(1100)						
US-10-292-798-833						
Alignment Scores:						
Pred. No.:	6.31e-107	Length:	1300			
Score:	978.50	Matches:	181			
Percent Similarity:	73.44%	Conservative:	54			
Best Local Similarity:	56.56%	Mismatches:	82			
Query Match:	56.96%	Indels:	3			
DB:	16	Gaps:	2			
US-10-081-775-2 (1-329) x US-10-292-798-833 (1-1300)						
QY	1	MetSerSerThrLeuGlyHisAsnMetGluSerProHisHisThrAspValAspProSer	20			
Db	135	CTAACCACTTACTTGGCCATGAACTCATTAACCATACCATCAG-----AACCCNACC	188			
QY	21	ValPhePheLeuLeuGlyIleProGlyLeuGluInPheHisLeuTrpLeuSerLeuPro	40			
Db	189	TCCTTTCTGCTCATGGGAATTCAGGCCCGGAGGCATCCACATTTTGGATTGCTTTTCCC	248			
QY	41	ValCysGlyLeuGlyThrAlaThrIleValGlyAsnIleThrIleLeuValValAla	60			
Db	249	TTCTGCTCCATGATGCTGGCAGTGTGGAAACATGGTGGTCTAGTGGTACAT	308			
QY	61	ThrGluProValLeuHisLysProValTyrLeuPheLeuCysMetLeuSerThrIleAsp	80			
Db	309	TCAGAGCTGTATTGCACCCACCATGATACCTGTTCTCTGTCATGCTATCCACATGGAC	368			
QY	81	LeuAlaAlaSerValSerThrValProLysLeuLeuAlaIlePheTrpCysGlyAlaGly	100			
Db	369	CTGCTCTCTGCTACCTCCATGTCGCCAAGCTCTTGCACTTTTGGGCAAGGATGCT	428			
QY	101	HisIleSerAlaSerAlaCysLeuAlaHisMetPhePheIleHisAlaPheCysMetMet	120			
Db	429	GAGATCAACTTGGGGCTGTGCTGCCAGATGTTCTTATCCATGCTTCTCAGCTGTA	488			
QY	121	GluSerThrValLeuLeuAlaMetAlaPheAspArgTyrValAlaIleCysHisProLeu	140			

Alignment Scores: 1.03e-103 Length: 1242
Pred. No.: 951.50 Matches: 180
Score:

Percent Similarity:	73.12%	Conservative:	54
Best Local Similarity:	56.25%	Mismatches:	83
Query Match:	55.38%	Indels:	4
DB:	15	Gaps:	2
US-10-081-775-2 (1-329) x US-10-017-161-959 (1-1242)			
Qy	1	MetSerSerThrLeuGlyHisAsnMetGluSerProHisHisThrAspValAspProSer	20
Db	78	CTAACCACTTACTTGGCCATGAACCTCATAAACCATACCATCAG-----AACCCCAACC	131
Qy	21	ValPhePheLeuGlyIleProGlyLeuGluGlnPheHisLeuThrLeuSerLeuPro	40
Db	132	TCCTTTCTGCTCATGGGAATTCAGGCCCGGAGGATCCCACTTTTGGATTGCTTTCC	191
Qy	41	ValCysGlyLeuGlyThrAlaThrIleValGlyAsnIleThrIleLeuValValAla	60
Db	192	TTCTGCTCCATGATGCCCTGGCAGTGTGGGAAACATGGTGGTCTGCTAGTGTACAT	251
Qy	61	ThrGluProValLeuHisGlyProValThrLeuPheLeuCysMetLeuSerThrIleAsp	80
Db	252	TCAGAGCTGTATTGACACAGCCCATGTACCTGCTTCATGCTATCCACCATTCAC	311
Qy	81	LeuAlaAlaSerValSerThrValProLysLeuLeuAlaIlePheThrCysGlyAlaGly	100
Db	312	CTGGTCTCTGCACCTCCACTGTGCCAAGCTCTTGCACCTTTTGGCAAGGATGCT	371
Qy	101	HisIleSerAlaSerAlaCysLeuAlaHisMetPhePheIleHisAlaPheCysMetMet	120
Db	372	GAGATCAACTTTGGGGCTGTGTGCCAGATGTTCTTATCCATGGCTTCTCAGCTGTA	431
Qy	121	GluSerThrValLeuAlaMetAlaPheAspArgThrValAlaIleCysHisProLeu	140
Db	432	GAATCTGTATACTCTAGCATGGCTTTGACCCCTACTTACGCTATTTGGGCTCTG	491
Qy	141	ArgThrAlaThrIleLeuThrAspThrIleAlaHisIleGlyValAlaAlaValVal	160
Db	492	CACATATGGTCTATCTCTCCAGAGCTGTAGGCAAGCTGGGGCTGCAGC-GTGCTT	550
Qy	161	ArgGlySerLeuLeuMetLeuProCysProPheIleGlyArgLeuAsnPheCysGln	180
Db	551	CGTGTGGGACTCATACCCCACTCACTCTTACTGGCAAGACTGAGTACTGTC---	607
Qy	181	SerHisValIleLeuHisThrCysGluHisMetAlaValValIleLeuAlaCysGly	200
Db	608	AGTCAGTGGTGGCCCACTCTCTACTGTGAACACATGGCTGTGAAGCTGGCTTGTGA	667
Qy	201	AspThrArgProAsnArgValThrGlyLeuThrAlaAlaLeuValIleGlyValAsp	220
Db	668	GGAAACACAGCCMAACAATCTATGGCATCTGTGCACACTGGTGGTGGCACTGAC	727
Qy	221	LeuPheCysIleGlyLeuSerThrAlaLeuIleAlaGlnAlaValLeuArgLeuSerSer	240
Db	728	TCCATCTGTATTGCTCTCTCTATGCACTCATCTCCGAGCTGTGTAGGTCTTCTCCTC	787
Qy	241	HisGluAlaArgSerIlyAlaLeuGlyThrCysGlySerHisValCysValIleLeuIle	260
Db	788	AAGGAGGCAAGGCTAAGACCTTTGGCACTTGTGGCTCCCACTGGGTGTCTACTTCTC	847
Qy	261	SerThrThrProAlaLeuPheSerPheThrHisArgPheGlyHisHisValProVal	280
Db	848	TTCTACACACAGGACTCTCTCTCTCTACACACAGCGGTGGCCAGCAGCTGCCCGG	907
Qy	281	HisIleHisIleLeuAlaAsnValThrLeuLeuProProAlaLeuAsnProVal	300
Db	908	CACATCCATCCTTCTAGCTACCTCTACCTGGTTGTGCCACCCCACTGTCAACCCCATC	967
Qy	301	ValThrGlyValIlyThrLysGlnIleAtgLysArgValValArgValPheGlnSerGly	320
Db	968	ATCTATGGCATGAAGACCAACAGATCTGGATGGGGCCCTCCGGCTTCTGAGTGGGSC	1027
RESULT 12			
US-10-343-650A-315			

; Sequence 315, Application US/10343650A			
; Publication No. US20040067499A1			
; GENERAL INFORMATION:			
; APPLICANT: HAGA, TATSUYA			
; TITLE OF INVENTION: NOVEL G-PROTEIN COUPLED RECEPTOR			
; FILE REFERENCE: 31671-186347			
; CURRENT APPLICATION NUMBER: US/10/343,650A			
; CURRENT FILING DATE: 2003-07-21			
; PRIOR APPLICATION NUMBER: JP 2000/237818			
; PRIOR FILING DATE: 2000-08-04			
; PRIOR APPLICATION NUMBER: JP 2001/34434			
; PRIOR FILING DATE: 2001-02-13			
; NUMBER OF SEQ ID NOS: 694			
; SOFTWARE: PatentIn Ver. 2.1			
; SEQ ID NO 315			
; LENGTH: 990			
; TYPE: DNA			
; ORGANISM: Homo sapiens			
; FEATURE:			
; NAME/KEY: CDS			
; LOCATION: (1)..(990)			
US-10-343-650A-315			
Alignment Scores:			
Pred. No.:	1.81e-101	Length:	990
Score:	931.50	Matches:	179
Percent Similarity:	73.91%	Conservative:	42
Best Local Similarity:	59.87%	Mismatches:	77
Query Match:	54.22%	Indels:	1
DB:	13	Gaps:	1
US-10-081-775-2 (1-329) x US-10-343-650A-315 (1-990)			
Qy	19	ProSerValPhePheLeuGlyIleProGlyLeuGluGlnPheHisLeuThrLeuSer	38
Db	82	CTCTCT---TTCTCTGTAGGGATTCAGTTTAGAGGAAGCACCACACTGGATTGCA	138
Qy	39	LeuProValCysGlyLeuGlyThrAlaThrIleValGlyAsnIleThrIleLeuValVal	58
Db	139	CTGCCCTGGGCATCTTTTACCTCTCTTTAGTGGCAATGTTTACCATTCTCTTCATC	198
Qy	59	ValAlaThrGluProValLeuHisIlyProValThrLeuPheLeuCysMetLeuSerThr	78
Db	199	ATCTGGATGGACCCATCTCTGCACCAATCTATGTACCTCTTCTCTCCATGCTAGTCC	258
Qy	79	IleAspLeuAlaAlaSerValSerThrValProLysLeuLeuAlaIlePheThrCysGly	98
Db	259	ATCGACCTGGTCTGGCTCTCTCCACTGCACCCAAAGCCCTTGCAGTGTCTCTGTTCT	318
Qy	99	AlaGlyHisIleSerAlaSerAlaCysLeuAlaHisMetPheIleHisAlaPheCys	118
Db	319	GCCACAGGATGGGTACATCGTCTGCTGATCCAGATGTTCTTCATCCATCACTTCTCC	378
Qy	119	MetMetGluSerThrValLeuLeuAlaMetAlaPheAspArgThrValAlaIleCysHis	138
Db	379	TCCATGGAGTCAGGGGTACTTGTGGCCATGGCTCTGGATCGCTATGTAGCCATTTGTC	438
Qy	139	ProLeuArgThrAlaThrIleLeuThrAspThrIleIleAlaHisIleGlyValAlaAla	158
Db	439	CCCTTGGACCATTCACCAATCTCGATCCAGGGGTCATAGGCGCATCGGAATGGTGTG	498
Qy	159	ValValArgGlySerLeuLeuMetLeuProCysProPheIleGlyArgLeuAsnPhe	178
Db	499	CTGTGTAGGGGATTACTACTCTCTTATCCCTTCCCATTTTGTGGAACTTATCTCTTC	558
Qy	179	CysGlnSerHisValIleLeuHisThrThrCysGluHisMetAlaValValIlyLeuAla	198
Db	559	TGCCAAGCCACCATCATAGGCCATGCCTATTGTGAACATATGCTGTGTGAACATGCC	618
Qy	199	CysGlyAspThrArgProAsnArgValThrGlyLeuThrAlaAlaLeuValIleGly	218
Db	619	TGCTCAGAAACACAGTCAATCGAGCTTATGGCTGACTATGGCTTGTGTGATTGGG	678

```

US-10-387-629-157

Alignment Scores:
Pred. No.:      1,81e-101      Length:      990
Score:          331.50         Matches:    179
Percent Similarity: 73.91%      Conservative: 42
Best Local Similarity: 59.87%    Mismatches: 77
Query Match:      54.22%       Indels:     1
DB:              16           Gaps:       1

US-10-081-775-2 (1-329) x US-10-387-629-157 (1-990)

QY      19 ProSerValPhePheLeuLeuGlyIleProGlyLeuGluInPheHisLeuTrpLeuSer 138
Ddb      82 CTTTTCT--TTTCTCTGGTAGGGATTCCAGGTTTAGAGGAAAGCCAGCAGCTGGATTGCA 138
QY      39 LeuProValCysGlyLeuGlyThrAlaThrIleValGlyAsnIleThrIleLeuValVal 58
Ddb     139 CTGCCCCCTGGGCATCCCTTACTCTCTGTCTTTAGTGGCAATGTTACCATTTCTTTCATC 198
QY     59 ValAlaThrGluProValLeuHisIleValProValTyrLeuPheLeuCysMetLeuSerThr 78
Ddb     199 ATCTGGATGGACCCATCTCTGCACCAATCTATGTACCTTCTCTGTCCATGCTAGCTGCC 258
QY     79 IleAspLeuAlaAlaSerValSerThrValProIlyLeuLeuAlaIlePheTrpCysGly 98
Ddb     259 ATCGACCTGGTTCGTGGCTCTCTCCACGTGCCACCAAGCCCTTCGAGTGCTCTCTGGTTCA 318
QY     99 AlaGlyHisIleSerAlaSerAlaCysLeuAlaHisMetPhePheIleHisAlaPheCys 118
Ddb     319 GCCCAGCAGATTGGGTACATCGTCTGCTGATCCAGATGTTCTTCATCCATGATTTTCTCC 378

```

```
; PRIOR APPLICATION NUMBER: US 60/280,597
; PRIOR FILING DATE: 2001-03-30
; PRIOR APPLICATION NUMBER: US 60/281,107
; PRIOR FILING DATE: 2001-04-02
; PRIOR APPLICATION NUMBER: US 60/282,121
; PRIOR FILING DATE: 2001-04-06
; NUMBER OF SEQ ID NOS: 96
; SOFTWARE: PERL Program
; SEQ ID NO 93
; LENGTH: 1114
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; OTHER INFORMATION: Incyte ID No: 7475226CB1
US-10-467-252-93

Alignment Scores:
Pred. No.: 2,18e-101 Length: 1114
Score: 931.50 Matches: 179
Percent Similarity: 73.91% Conservative: 42
Best Local Similarity: 59.87% Mismatches: 77
Query Match: 54.22% Indels: 1
DB: 17 Gaps: 1

US-10-081-775-2 (1-329) x US-10-467-252-93 (1-1114)
QY 19 ProSerValPhePheLeuLeuGlyLeuProGlyLeuGluGlnPheHisLeuTrpLeuSer 38
Db 106 CTTCTCTTTCTCTGGTAGGATTCAGGTTTATAGGAAAGCCAGCAGCTGGATTGCA 162
QY 39 LeuProValCysGlyLeuGlyThrAlaThrIleValGlyAsnIleThrIleValVal 58
Db 163 CTGCCCTGGGATCTTACCTCTCTTGTAGTGGCAATGTTACCAATCTCTTCATC 222
QY 59 ValAlaThrGluProValLeuHisLysProValTyrLeuPheLeuCysMetLeuSerThr 78
Db 223 ATCTGGAGTGGCCATCTCTTGCCCAATCTATGACCTCTCTCTGTCATGCTAGTGC 282
QY 79 IleAspLeuAlaAlaSerValSerThrValProLysLeuLeuAlaIlePheTrpCysGly 98
Db 283 ATCGACCTGGTTCTGGCTCTCCCTCCACTGCACCCAAAGCCCTTGCAGTCTCTCTG 342
QY 99 AlaGlyHisIleSerAlaSerAlaCysLeuAlaHisMetPhePheIleHisAlaPheCys 118
Db 343 GCCCAGCAGATTGGGTACATCTGCTGCTGATCCAGATGTTCTTCATCCATGCAATCTCC 402
QY 119 MetMetGluSerThrValLeuLeuAlaMetAlaPheAspArgTyrValAlaIleCysHis 138
Db 403 TCCATGGAGTCAGGGGTACTTGTGGCCATGGCTCTGGATCGCTATGATGCCATTGTCTAC 462
QY 139 ProLeuArgTyrAlaThrIleLeuThrAspThrIleIleAlaHisIleGlyValAlaAla 158
Db 463 CCCTTGACCAATCCACAATCTGCATCCAGGGGTCTATAGGGCGCATAGGAAATGGTGTG 522
QY 159 ValValArgGlySerLeuLeuMetLeuProCysProPhePheIleGlyArgLeuAsnPhe 178
Db 523 CTGGTGGAGGGATTACTATCTCTTATCCCTTATCCCTTATCCCTTATGGGAACACTTATCTTC 582
QY 179 CysGlnSerHisValIleLeuHisThrTyrCysGluHisMetAlaValValLysLeuAla 198
Db 583 TGCCAGCCACCACATCATAGCCATGCTATTGTGAACATATGCTGTGTGAACATTGCC 642
QY 199 CysGlyAspThrArgProAsnArgValTyrGlyLeuThrAlaAlaLeuValIleGly 218
Db 643 TGCTCAAAACACACATCAATCAGCTTATGGGCTGACTATGGCTTGTGTGATGGG 702
QY 219 ValAspLeuPheCysIleGlyLeuSerTyrAlaLeuIleAlaGlnAlaValLeuArgLeu 238
Db 703 CTGATGTTCTGGCCATGTGTCTTATGCCCCACATCTCCAGGAGTGTCTGAAGTTA 762
QY 239 SerSerHisGluAlaArgSerLysAlaLeuGlyThrCysGlySerHisValCysValIle 258
Db 762 CTGATGTTCTGGCCATGTGTCTTATGCCCCACATCTCCAGGAGTGTCTGAAGTTA 762
Db 239 SerSerHisGluAlaArgSerLysAlaLeuGlyThrCysGlySerHisValCysValIle 258
Db 762 CTGATGTTCTGGCCATGTGTCTTATGCCCCACATCTCCAGGAGTGTCTGAAGTTA 762

; PRIOR APPLICATION NUMBER: US 60/280,597
; PRIOR FILING DATE: 2001-03-30
; PRIOR APPLICATION NUMBER: US 60/281,107
; PRIOR FILING DATE: 2001-04-02
; PRIOR APPLICATION NUMBER: US 60/282,121
; PRIOR FILING DATE: 2001-04-06
; NUMBER OF SEQ ID NOS: 96
; SOFTWARE: PERL Program
; SEQ ID NO 93
; LENGTH: 1114
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; OTHER INFORMATION: Incyte ID No: 7475226CB1
US-10-467-252-93

Alignment Scores:
Pred. No.: 2,18e-101 Length: 1114
Score: 931.50 Matches: 179
Percent Similarity: 73.91% Conservative: 42
Best Local Similarity: 59.87% Mismatches: 77
Query Match: 54.22% Indels: 1
DB: 17 Gaps: 1

US-10-081-775-2 (1-329) x US-10-467-252-93 (1-1114)
QY 19 ProSerValPhePheLeuLeuGlyLeuProGlyLeuGluGlnPheHisLeuTrpLeuSer 38
Db 106 CTTCTCTTTCTCTGGTAGGATTCAGGTTTATAGGAAAGCCAGCAGCTGGATTGCA 162
QY 39 LeuProValCysGlyLeuGlyThrAlaThrIleValGlyAsnIleThrIleValVal 58
Db 163 CTGCCCTGGGATCTTACCTCTCTTGTAGTGGCAATGTTACCAATCTCTTCATC 222
QY 59 ValAlaThrGluProValLeuHisLysProValTyrLeuPheLeuCysMetLeuSerThr 78
Db 223 ATCTGGAGTGGCCATCTCTTGCCCAATCTATGACCTCTCTCTGTCATGCTAGTGC 282
QY 79 IleAspLeuAlaAlaSerValSerThrValProLysLeuLeuAlaIlePheTrpCysGly 98
Db 283 ATCGACCTGGTTCTGGCTCTCCCTCCACTGCACCCAAAGCCCTTGCAGTCTCTCTG 342
QY 99 AlaGlyHisIleSerAlaSerAlaCysLeuAlaHisMetPhePheIleHisAlaPheCys 118
Db 343 GCCCAGCAGATTGGGTACATCTGCTGCTGATCCAGATGTTCTTCATCCATGCAATCTCC 402
QY 119 MetMetGluSerThrValLeuLeuAlaMetAlaPheAspArgTyrValAlaIleCysHis 138
Db 403 TCCATGGAGTCAGGGGTACTTGTGGCCATGGCTCTGGATCGCTATGATGCCATTGTCTAC 462
QY 139 ProLeuArgTyrAlaThrIleLeuThrAspThrIleIleAlaHisIleGlyValAlaAla 158
Db 463 CCCTTGACCAATCCACAATCTGCATCCAGGGGTCTATAGGGCGCATAGGAAATGGTGTG 522
QY 159 ValValArgGlySerLeuLeuMetLeuProCysProPhePheIleGlyArgLeuAsnPhe 178
Db 523 CTGGTGGAGGGATTACTATCTCTTATCCCTTATCCCTTATCCCTTATGGGAACACTTATCTTC 582
QY 179 CysGlnSerHisValIleLeuHisThrTyrCysGluHisMetAlaValValLysLeuAla 198
Db 583 TGCCAGCCACCACATCATAGCCATGCTATTGTGAACATATGCTGTGTGAACATTGCC 642
QY 199 CysGlyAspThrArgProAsnArgValTyrGlyLeuThrAlaAlaLeuValIleGly 218
Db 643 TGCTCAAAACACACATCAATCAGCTTATGGGCTGACTATGGCTTGTGTGATGGG 702
QY 219 ValAspLeuPheCysIleGlyLeuSerTyrAlaLeuIleAlaGlnAlaValLeuArgLeu 238
Db 703 CTGATGTTCTGGCCATGTGTCTTATGCCCCACATCTCCAGGAGTGTCTGAAGTTA 762
QY 239 SerSerHisGluAlaArgSerLysAlaLeuGlyThrCysGlySerHisValCysValIle 258
Db 762 CTGATGTTCTGGCCATGTGTCTTATGCCCCACATCTCCAGGAGTGTCTGAAGTTA 762
Db 239 SerSerHisGluAlaArgSerLysAlaLeuGlyThrCysGlySerHisValCysValIle 258
Db 762 CTGATGTTCTGGCCATGTGTCTTATGCCCCACATCTCCAGGAGTGTCTGAAGTTA 762

; Sequence 945, Application US/10017161
; Publication No. US20030143668A1
; GENERAL INFORMATION:
; APPLICANT: SUMA, MAKIKO
; APPLICANT: ASAI, KIYOSHI
; APPLICANT: AKIYAMA, YUTAKA
; APPLICANT: ABURATANI, HIROYUKI
; TITLE OF INVENTION: NOVEL G PROTEIN-COUPLED RECEPTORS
; FILE REFERENCE: 084335/0152
; CURRENT APPLICATION NUMBER: US/10/017,161
; CURRENT FILING DATE: 2002-12-18
; PRIOR APPLICATION NUMBER: JP 2001/246789
; PRIOR FILING DATE: 2001-06-18
; NUMBER OF SEQ ID NOS: 2430
; SOFTWARE: Patent In Ver. 2.1
; SEQ ID NO 945
; LENGTH: 1390
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: source
; LOCATION: (1)..(1390)
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (201)..(1190)
; FEATURE:
; NAME/KEY: modified base
; LOCATION: (1)..(26)
; OTHER INFORMATION: a, t, c, g, unknown or other
US-10-017-161-945

Alignment Scores:
Pred. No.: 3,07e-101 Length: 1390
Score: 931.50 Matches: 179
Percent Similarity: 73.91% Conservative: 42
Best Local Similarity: 59.87% Mismatches: 77
Query Match: 54.22% Indels: 1
DB: 15 Gaps: 1

US-10-081-775-2 (1-329) x US-10-017-161-945 (1-1390)
QY 19 ProSerValPhePheLeuLeuGlyLeuProGlyLeuGluGlnPheHisLeuTrpLeuSer 38
Db 282 CTTCTCTTTCTCTGGTAGGATTCAGGTTTATAGGAAAGCCAGCAGCTGGATTGCA 338
QY 39 LeuProValCysGlyLeuGlyThrAlaThrIleValGlyAsnIleThrIleValVal 58
Db 339 CTGCCCTGGGATCTTACCTCTCTTGTAGTGGCAATGTTACCAATCTCTTCATC 398
QY 59 ValAlaThrGluProValLeuHisLysProValTyrLeuPheLeuCysMetLeuSerThr 78
Db 399 ATCTGGATGACCCATCTTCACCAATCTATGATCTCTCTCTGTCATGCTAGTGC 458
QY 79 IleAspLeuAlaAlaSerValSerThrValProLysLeuLeuAlaIlePheTrpCysGly 98
Db 459 ATCGACCTGGTTCTGGCTCTCTCCACTGCACCCAAAGCCCTTGCAGTGTCTCTGTTTCA 518
```

QY 99 AlaGlyHisIleSerAlaSerAlaCysLeuAlaHisMetPhePheIleHisAlaPheCys 118
Db |||||
519 GCCACGAGATTGGGTACATCTCTGCTGATCCAGATGTTCTTCATCCATGCAITCTCC 578
QY 119 MetMetGluSerThrValLeuLeuAlaMetAlaPheAspArgTyrValAlaIleCysHis 138
Db |||||
579 TCCATGGAGTACAGGGTACTTGTGGCCATGGCTCTGGATCGTATGTAGCCATTGTGCAC 638
QY 139 ProLeuArgTyrAlaThrIleLeuThrAspThrIleIleAlaHisIleGlyValAlaAla 158
Db |||||
639 CCCCTGCACCATTCACAACTCTGATCCAGGGGTCATAGGCGCATCGGAATGGTGGTG 698
QY 159 ValValArgGlySerLeuLeuMetLeuProCysProPhePheIleGlyArgLeuAsnPhe 178
Db |||||
699 CTGGTGAGGGATTACTACTCTTATCCCTTCCCATTTTGTGGGAACACTTATCTTC 758
QY 179 CysGlnSerHisValIleLeuHisThrTyrCysGluHisMetAlaValValIysLeuAla 198
Db |||||
759 TGCCAAGCCACCATCATAGGCCATGCTATTGTGAACATATGCTGTGTGAACCTTGCC 818
QY 199 CysGlyAspThrArgProAsnArgValTyrGlyLeuThrAlaAlaLeuLeuValIleGly 218
Db |||||
819 TGCTCAGAACACACAGTCAATCGAGCTTATGGGCTGACTATGGCCTTGTGTGATTGGG 878
QY 219 ValAspLeuPheCysIleGlyLeuSerTyrAlaLeuIleAlaGlnAlaValLeuArgLeu 238
Db |||||
879 CTGGATGTTCTGGCCATTGGTGTTCCTATGCCACATCCTCCAGCGAGTCTGAAGGTA 938
QY 239 SerSerHisGluAlaArgSerIysAlaLeuGlyThrCysGlySerHisValCysValIle 258
Db |||||
939 CCAGGGAGTGAAGCCCGACTTAAGGCGTTTAGCACATGTGGCTCTCATATTGTGTCTATC 998
QY 259 LeuIleSerTyrThrProAlaLeuPheSerPhePheThrHisArgPheGlyHisVal 278
Db |||||
999 CTGTCCTCTATGTCCTCGGAATTTCTCTCTCACTACCGCTTTGGTCAATCATGTA 1058
QY 279 ProValHisIleHisIleLeuLeuAlaAsnValTyrLeuLeuLeuProProAlaLeuAsn 298
Db |||||
1059 CCCCATCATGTCCATGTTCTTCTGGCCACACGGTATCTCTCATGCCACCTGGCGTCMAAT 1118
QY 299 ProValValTyrGlyValIysGlnIleArgLysArgValValArgValPhe 317
Db |||||
1119 CCTCTGTCTATGGAGTGAGACTCAGCAGATCCGCCAGCGAGTGTCTCAGAGTGTTT 1175

Search completed: August 27, 2004, 21:07:04
Job time : 556 secs

This Page Blank (b)(7)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: August 27, 2004, 18:30:23 ; Search time 90 Seconds

(without alignments)

2028.654 Million cell updates/sec

Title: US-10-081-775-2

Perfect score: 1718

Sequence: 1 MSSTLGHMESPHTVDPS.....RKRVRVFSQGGMIKASE 329

Scoring table: BLOSUM62

Xgapop 10.0 , Xgapext 0.5

Ygapop 10.0 , Ygapext 0.5

Fgapop 6.0 , Fgapext 7.0

Delop 6.0 , Delext 7.0

Searched: 682709 seqs, 277475446 residues

Total number of hits satisfying chosen parameters: 1365418

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+p2n.model -DEV=xlh

-Q=/cgn2_1/USPTO.spool/US10081775/runat_20082004_170214_11153/app_query.fasta_1.519

-DB=issued Patents NA -QFMT=fastap -SUFFIX=rni -MINMATCH=0.1 -LOOPCL=0

-LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdd

-LIST=45 -DOCALLIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15

-MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZ=500 -MINLEN=0 -MAXLEN=2000000000

-USER=US10081775@cgn_1_69@runat_20082004_170214_11153 -NCPU=6 -ICPU=3

-NO_MMAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSBLOCK=100 -LONGLOG

-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6

-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

Issued Patents NA.*

1: /cgn2_6/ptodata/2/ina/5A.COMB.seq.*

2: /cgn2_6/ptodata/2/ina/5B.COMB.seq.*

3: /cgn2_6/ptodata/2/ina/6A.COMB.seq.*

4: /cgn2_6/ptodata/2/ina/6B.COMB.seq.*

5: /cgn2_6/ptodata/2/ina/PCTUS.COMB.seq.*

6: /cgn2_6/ptodata/2/ina/backfiles1.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	867.5	50.5	1539	US-09-668-680-13	Sequence 13, Appl
2	861	50.1	963	US-09-439-313-526	Sequence 526, App
3	861	50.1	963	US-09-636-215-526	Sequence 526, App
4	861	50.1	963	US-09-685-166A-526	Sequence 526, App
5	856	49.8	1474	US-08-465-980-1	Sequence 1, Appli
6	856	49.8	1474	US-09-053-303-1	Sequence 1, Appli
7	856	49.8	1474	US-09-339-115-1	Sequence 1, Appli
8	856	49.8	1474	PCT-US95-07093-1	Sequence 1, Appli
9	424	24.7	966	US-08-748-506-5	Sequence 5, Appli
10	423.5	24.7	966	US-08-748-506-7	Sequence 7, Appli
11	416.5	24.2	966	US-08-748-506-6	Sequence 6, Appli
12	410.5	23.9	966	US-08-748-506-8	Sequence 8, Appli

13	410	23.9	1282	4	US-09-016-434-1413	Sequence 1413, Ap
14	410	23.9	1351	4	US-09-546-986A-5	Sequence 5, Appli
15	410	23.9	1351	4	US-09-524-730-5	Sequence 5, Appli
16	404.5	23.5	1438	4	US-09-016-434-1313	Sequence 1313, Ap
17	401	23.3	1035	4	US-09-546-986A-1	Sequence 1, Appli
18	401	23.3	1035	4	US-09-524-730-1	Sequence 1, Appli
19	398.5	23.2	1854	4	US-09-016-434-1312	Sequence 1312, Ap
20	396.5	23.1	1713	2	US-08-467-948A-1	Sequence 1, Appli
21	396.5	23.1	1713	3	US-08-467-947A-1	Sequence 1, Appli
22	390	22.7	1062	4	US-09-668-680-10	Sequence 10, Appl
23	387	22.5	945	4	US-09-016-434-1114	Sequence 1114, Ap
24	385	22.4	1990	4	US-09-016-434-1056	Sequence 1056, Ap
25	381.5	22.2	900	3	US-09-085-371-5	Sequence 5, Appli
26	381	22.2	1065	4	US-09-546-986A-7	Sequence 7, Appli
27	381	22.2	1065	4	US-09-524-730-7	Sequence 7, Appli
28	376.5	21.9	3459	4	US-09-016-434-1363	Sequence 1363, Ap
29	374	21.8	1290	2	US-08-827-291A-1	Sequence 1, Appli
30	373.5	21.7	952	4	US-09-016-434-1115	Sequence 1115, Ap
31	352	20.5	1080	4	US-09-668-680-9	Sequence 9, Appli
32	348.5	20.3	1297	4	US-09-668-680-11	Sequence 11, Appl
33	325	18.9	984	3	US-08-748-506-9	Sequence 9, Appli
34	320	18.6	951	4	US-09-465-901-47	Sequence 47, Appl
35	311.5	18.1	1411	4	US-09-546-986A-3	Sequence 3, Appli
36	311.5	18.1	1411	4	US-09-524-730-3	Sequence 3, Appli
37	310.5	18.1	666	4	US-09-465-901-33	Sequence 33, Appl
38	305	17.8	678	4	US-09-465-901-45	Sequence 45, Appl
39	299.5	17.4	675	4	US-09-465-901-43	Sequence 43, Appl
40	297	17.3	675	4	US-09-465-901-17	Sequence 17, Appl
41	293	17.1	669	4	US-09-465-901-37	Sequence 37, Appl
42	289	16.8	669	4	US-09-465-901-11	Sequence 11, Appl
43	287	16.7	669	4	US-09-465-901-35	Sequence 35, Appl
44	285	16.6	669	4	US-09-465-901-39	Sequence 39, Appl
45	283	16.5	669	4	US-09-465-901-15	Sequence 15, Appl

ALIGNMENTS

RESULT 1

US-09-668-680-13
; Sequence 13, Application US/09668680
; Patent No. 6436703

; GENERAL INFORMATION:

; APPLICANT: Tang, Y. Tom

; APPLICANT: Liu, Chenghua

; APPLICANT: Zhou, Ping

; APPLICANT: Asundi, Vinod

; APPLICANT: Zhang, Jie

; APPLICANT: Wang, Jian-Rui

; APPLICANT: Xue, Aidong J.

; APPLICANT: Xu, Chongjun

; APPLICANT: Drmanac, Radoje T.

; TITLE OF INVENTION: No. 6436703el Nucleic Acids and

; TITLE OF INVENTION: Polypeptides

; FILE REFERENCE: 790CIP2A

; CURRENT APPLICATION NUMBER: US/09/668,680

; PRIOR FILING DATE: 2000-09-22

; PRIOR APPLICATION NUMBER: 09/649,167

; PRIOR FILING DATE: 2000-08-23

; PRIOR APPLICATION NUMBER: 09/540,217

; PRIOR FILING DATE: 2000-03-31

; NUMBER OF SEQ ID NOS: 13

; SOFTWARE: pt_FL_genes Version 2.0

; SEQ ID NO 13

; LENGTH: 1539

; TYPE: DNA

; ORGANISM: Homo sapiens

; FEATURE:

; NAME/KEY: CDS

; LOCATION: (130)..(1539)

US-09-668-680-13

Alignment Scores:

Pred. No.: 1.04e-92 Length: 1539

```
Score: 867.50 Matches: 156
Percent Similarity: 69.70% Conservative: 51
Best Local Similarity: 52.53% Mismatches: 89
Query Match: 50.49% Indels: 1
DB: 4 Gaps: 1

US-10-081-775-2 (1-329) x US-09-658-680-13 (1-1539)

Qy 26 GlyIleProGlyLeuGluGlnPheHisLeuTrpLeuSerLeuProValCysGlyLeuGly 45
Db 637 GGGATCCAGGCTGGAGGCTGCCACTTCCTGGATTGCCATCCCTTTCTGTGGCCATGAT 696
Qy 46 ThrAlaThrIleValGlyAsnIleThrIleLeuValValAlaThrGluProValIleu 65
Db 697 CTTGTAGCACTGGTGGAAATGCTGCCCTCATCTGGCTATTGGCCATGGACAATGCTCTT 756
Qy 66 HisIysProValTyrLeuPheLeuCysMetLeuSerThrIleAspLeuAlaAlaSerVal 85
Db 757 CATGCACCTATGTACTCTTCTCTGCTTCTCTCACTCACAGACCTGGCTCTCAGTTCT 816
Qy 86 SerThrValProIysLeuLeuAlaIlePheTrpCysGlyAlaGlyHisIleSerAlaSer 105
Db 817 ACCACTGTGCCAAGATGCTGGCCATTTTGTGGCTCCATGCTGTGATTCCTTTGGT 876
Qy 106 AlaCysLeuAlaHisMetPhePheIleHisAlaPheCysMetMetGluSerThrValIleu 125
Db 877 GGATGCTGGCCAGATGTTTGTGTCATTTCTATCTATGCTCTGGAGTCTCTCGATCTA 936
Qy 126 LeuAlaMetAlaPheAspArgTyrValAlaIleCysHisProLeuArgTyrAlaThrIle 145
Db 937 CTTGCCATGGCTTTGATAGTATGGCTATCTGTAACCCATTAAGGTATACAAACCAT 996
Qy 146 LeuThrAspThrIleIleAlaHisIleGlyValAlaValAlaValArgGlySerLeuLeu 165
Db 997 CTCAACCATGCTGTATAGGAGCAATGGCTTTGTTGGCTATTCCGTAGTGTGGCTATT 1056
Qy 166 MetLeuProCysProPhePheIleGlyArgLeuAsnPheCysGlnSerHisValIleLeu 185
Db 1057 GTCTCCCTTCATCTTCTGCTGAGGCGATCCCTTACTGTGTGCTACCGTGTGATGACA 1116
Qy 186 HisThrTyrCysGluHisMetAlaValIysLeuAlaCysGlyAspThrArgProAsn 205
Db 1117 CACACATACTGTAGCATATGGCATCGCCGAGCTGGCTGTGCCAACATCACTGTCAAT 1176
Qy 206 ArgValTyrGlyLeuThrAlaAlaLeuValIleGlyValAspLeuPheCysIleGly 225
Db 1177 ATTCTATGGCTTAACCTGTGGCTCTGCTGGCCATGGGACTGGATTCATTTCTCATTTGCC 1236
Qy 226 LeuSerTyrAlaLeuIleAlaGlnAlaValLeuArgLeuSerSerHisGluAlaArgSer 245
Db 1237 ATTTCTATGGCTTTATCTCTCATGCTGCTTTTCCCTTCCATCTCATGATGCCAGCAC 1296
Qy 246 LysAlaLeuGlyThrCysGlySerHisValIleCysValIleLeuIleSerThrProAla 265
Db 1297 AAAGCTGTAGTACTGTGGCTCCACATGGCATCATCTGGTGTTCATCATCCCTGCC 1356
Qy 266 LeuPheSerPheThrHisArgPheGlyHisHis---ValProValHisIleHisIle 284
Db 1357 TTCTTCTCTTCTCACCCACCGCTTGGTGGTCCACCAAGGAGTCCCAAGCATGTGCACATC 1416
Qy 285 LeuLeuAlaAsnValTyrLeuLeuProProAlaLeuAsnProValValTyrGlyVal 304
Db 1417 TTTCTGGCTATCTATGCTGCTGGTCTCTCTGTACTCATCTCTATTTCTATGGAGCT 1476
Qy 305 LysThrIysGlnIleArgLysArgValValArgValPheGlnSerGlyGln 321
Db 1477 AGAACCAAGGAGATTCGGAGTCGACTTCTAAAACTGCTTCACTCCCTGGGGAAG 1527

RESULT 2
US-09-439-313-526
; Sequence 526, Application US/09439313
; Patent No. 6329505
; GENERAL INFORMATION:
```

```
APPLICANT: Xu, Jiangchun
APPLICANT: Dillon, Davin C.
APPLICANT: Mitcham, Jennifer L.
APPLICANT: Harlocker, Susan Louise
APPLICANT: Jiang Yuqi
APPLICANT: Reed, Steven G.
APPLICANT: Kalos, Michael
APPLICANT: Fanger, Gary
APPLICANT: Retter, Mark
APPLICANT: Solk, John
APPLICANT: Day, Craig
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THERAPY AND
FILE REFERENCE: 210121.427C9
CURRENT APPLICATION NUMBER: US/09/439,313
NUMBER OF SEQ ID NOS: 575
SOFTWARE: Fast-Seq for Windows Version 3.0
SEQ ID NO 526
LENGTH: 963
TYPE: DNA
ORGANISM: Homo sapiens
US-09-439-313-526

Alignment Scores:
Pred. No.: 2,95e-92 Length: 963
Score: 861.00 Matches: 161
Percent Similarity: 70.23% Conservative: 49
Best Local Similarity: 53.85% Mismatches: 89
Query Match: 50.12% Indels: 0
DB: 4 Gaps: 0

US-10-081-775-2 (1-329) x US-09-439-313-526 (1-963)

Qy 20 SerValPhePheLeuLeuGlyIleProGlyLeuGluGlnPheHisLeuTrpLeuSerLeu 39
Db 25 GCCACCTTTGTGCTTATTGTTATCCAGGATTAGAGAAAGCCCATTTCTGGTGGTTC 84
Qy 40 ProValCysGlyLeuGlyThrAlaThrIleValGlyAsnIleThrIleLeuValValVal 59
Db 85 CCCCTCTCTTCCATGTATGTAGTGGCAATGTTTGGAACTGTCATCGTGGTCTTCATCGTA 144
Qy 60 AlaThrGluProValLeuHisLysProValTyrLeuPheLeuCysMetLeuSerThrIle 79
Db 145 AGGACGAGAACGACCTTCGACCTCCGATGACCTCTTCTCTGATGCTTGACCCATT 204
Qy 80 AspLeuAlaAlaSerValSerThrValProIysLeuLeuAlaIlePheTrpCysGlyAla 99
Db 205 GACCTGGCTTATCCACATCCACCATCCATGCTAAGATCCTTGCCCTTTCTGCTTTGATTCC 264
Qy 100 GlyHisIleSerAlaSerAlaCysLeuAlaHisMetPhePheIleHisAlaPheCysMet 119
Db 265 CGAGAGATTAGCTTTGAGGCTGTCTTACCAGATGTTCTTTATTCATGCCCTCTCAGCC 324
Qy 120 MetGluSerThrValLeuLeuAlaMetAlaPheAspArgTyrValAlaIleCysHisPro 139
Db 325 ATTGAATCCACCATCTCTGCGCATGGCTTTGACCGTTATGTGGCATCTGCCACCCA 384
Qy 140 LeuArgTyrAlaThrIleLeuThrAspThrIleAlaHisIleGlyValAlaAlaVal 159
Db 385 CTGGCCATGTGCGATGCTCAACATACAGCCAGCCAGCATGGCATCTGCTGGCTGTG 444
Qy 160 ValArgGlySerLeuLeuMetLeuProCysProPhePheIleGlyArgLeuAsnPheCys 179
Db 445 GTCCGGGATTCCTCTTTTTCCTCCCTGCTCTGCTGATCAAGCGCTGGCTTCTGCT 504
Qy 180 GlnSerHisValIleLeuHisThrTyrCysGluHisMetAlaValValIysLeuAlaCys 199
Db 505 CACTCCAATGCTCTCGCATCTCTATTGTGTGCCAGGATGTAATGAAGTTGGCTTAT 564
Qy 200 GlyAspThrArgProAsnArgValTyrGlyLeuThrAlaAlaLeuLeuValIleGlyVal 219
Db 565 GCAGACACTTTGGCCCAATGTGTATGTGCTTACTGCACTTCTGCTGCTCATGGGCGTG 624
```


Qy	220	AspLeuPheCysIleGlyLeuSerTyrAlaLeuIleAlaGlnAlaValLeuArgLeuSer	239
Db	625	GACGTAATGTTTCATCTCCTGTGCTTATTTCTGATAAATACGACGGTCTGCAACTGCCT	684
Qy	240	SerHisGluAlaArgSerLysAlaLeuGlyThrCysGlySerHisValCysValIleLeu	259
Db	685	TCCAAGTCAGACGGCGCAAGCCCTTGGAACTGTGTGCACATGTGTTGGTGATCTC	744
Qy	260	IleSerTyrThrProAlaLeuPheSerPheThrHisArgPheGlyHisValPro	279
Db	745	GCCCTTCATGTGCCACTTATTGGCCCTCTCAGTGTGTACACCGCTTTGGAAACAGCCTTCAT	804
Qy	280	ValHisIleHisIleLeuLeuAlaAsnValTyrLeuLeuLeuProProAlaLeuAsnPro	299
Db	805	CCCATTGGCGGTGTGTGTCATGGGTGCATCTACCTGCTGCTGCTCTCTCATCAATCC	864
Qy	300	ValValTyrGlyValLysThrLysGlnIleArgLysArgValValArgValPheGln	318
Db	865	ATCATCTATGTGTGCAAAACCAAAACAGATCAAGAACACGGGTCTGGCATGTGTTCAAG	921

RESULT 3

US-09-636-215-526
 ; Sequence 526, Application US/09636215
 ; Patent No. 6620922
 ; GENERAL INFORMATION:
 ; APPLICANT: Xu, Jiangchun
 ; APPLICANT: Dillon, Davin C.
 ; APPLICANT: Mitcham, Jennifer L.
 ; APPLICANT: Harlocker, Susan L.
 ; APPLICANT: Jiang, Yuqi
 ; APPLICANT: Henderson, Robert A.
 ; APPLICANT: Kalos, Michael D.
 ; APPLICANT: Fanger, Gary R.
 ; APPLICANT: Retter, Marc W.
 ; APPLICANT: Stolk, John A.
 ; APPLICANT: Day, Craig H.
 ; APPLICANT: Vedvick, Thomas S.
 ; APPLICANT: Carter, Darrick
 ; APPLICANT: Li, Samuel
 ; APPLICANT: Wang, Aijun
 ; APPLICANT: Skeiky, Yasir A.W.
 ; APPLICANT: Hepler, William
 ; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
 ; TITLE OF INVENTION: DIAGNOSIS OF PROSTATE CANCER
 ; FILE REFERENCE: 210121.42717C17
 ; CURRENT APPLICATION NUMBER: US/09/636,215
 ; CURRENT FILING DATE: 2000-08-10
 ; NUMBER OF SEQ ID NOS: 852
 ; SOFTWARE: FastSeq for Windows Version 3.0
 ; SEQ ID NO 526
 ; LENGTH: 963
 ; TYPE: DNA
 ; ORGANISM: Homo sapiens
 US-09-636-215-526

```

Alignment Scores:
Pred. No.:          2,95e-92          Length:          963
Score:             861.00             Matches:         161
Percent Similarity: 70.23%            Conservative:    49
Best Local Similarity: 53.85%          Mismatches:     89
Query Match:       50.12%              Indels:         0
DB:                4                  Gaps:          0

US-10-081-775-2 (1-329) x US-09-636-215-526 (1-963)

Qy      20  SerValPhePheLeuGlyIleProGlyLeuGluGlnPheHisLeuTrpLeuSerLeu 39
      ::  :::: :::: :::: :::: :::: :::: :::: :::: :::: ::::
Db      25  GCCACCTTGTGCTTATTTGGTATCCAGATTAGAGAAAGCCCATTTCTGGTTGCCTTC 84

Qy      40  ProValCysGlyLeuGlyThrAlaThrIleValGlyAsnIleThrIleLeuValVal 59
      :::: :::: :::: :::: :::: :::: :::: :::: :::: ::::
Db      85  CCCCTCTTCCATGTATGTATGTGCAATGTTTCGAAATGTCATCGTGGTCTTCACTCGTA 144

```

60 AlaThrGluProValLeuHisHisProValTyrLeuPheLeuCysMetLeuSerThr 79
145 AGGACGGAAACGACGCTGCACGCTCGAATGATACCTCTTCTCTGCATGCTTGCAGCCATT 204
80 AspLeuAlaAlaSerValSerThrValProLysLeuLeuAlaAlaPheTrpCysGlyAla 99
205 GACCTGGCGCTTATCCACATCCACCATGCCTAAGATCCTTGCCTTCTGGTTTGATTCC 264
100 GlyHisIleSerAlaSerAlaCysLeuAlaHisMetPhePheIleHisAlaPheCysMet 119
265 CGAGATATAGCTTTGAGGCTCTCTTACCAGATGTTCTTTATTCATGCCTCTCAGCC 324
120 MetGluSerThrValLeuLeuAlaMetAlaPheAspArgTyrValAlaIleCysHisPro 139
325 ATTGAATCCACATCCTCTGGCCATGGCTTTGACCGTTATGTGGCCATCTGCCACCCA 384
140 LeuArgTyrAlaThrIleLeuThrAspThrIleIleAlaHisIleGlyValAlaAlaVal 159
385 CTGGCCATGCTGCAGTGTCTCAACAATACAGTAACAGCCACAGATTGGCATCGTGGCTGTG 444
160 ValArgGlySerLeuLeuMetLeuProCysProPhePheIleGlyArgLeuAsnPheCys 179
445 GTCCGGAGATCCCTCTTTTCCCATCGCTCTGCTGCTCATCAAGCGGTGGCTTCTGC 504
180 GlnSerHisValIleLeuHisThrTyrCysGluHisMetAlaValValLysLeuAlaCys 199
505 CACTCCAATGTCTCTGCACCTCTCTATTGTGTCCACAGGATGTAATGAAGTTGGCTAT 564
200 GlyAspThrArgProAsnArgValTyrGlyLeuThrAlaAlaLeuLeuValIleGlyVal 219
565 GCAGACACTTGGCCAAATGTGTATATGCTCTTCTGCCATTTCTGCTGTGTATGGCGGTG 624
220 AspLeuPheCysIleGlyLeuSerTyrAlaLeuIleAlaGlnAlaValLeuArgLeuSer 239
625 GACGTAAATGTTCACTCTCTGCTGCTTATTCTTGATAATACGAACGGTCTCTGCAACTGCCT 684
240 SerHisGluAlaArgSerLysAlaLeuGlyThrCysGlySerHisValCysValIleLeu 259
685 TCCAAGTCAGACGGCGGCGCCCTTGGAACTGTGTGTACACATGGTGTGTACTCTC 744
260 IleSerTyrThrProAlaLeuPheSerPhePheThrHisArgPheGlyHisIleValPro 279
745 GCCTTCTATGTGCCACTATTGGCTCTCAGTTGTACACCGCTTTGGAAACAGCCTTCAT 804
280 ValHisIleHisIleLeuLeuAlaAsnValTyrLeuLeuLeuProProAlaLeuAsnPro 299
805 CCCATTGTGCGTGTGTTCATGGGTGACATCTACCTGCTGCTGCTCTCTCATCAATCCC 864
300 ValValTyrGlyValLysThrLysGlnIleArgLysArgValValArgValPheGln 318
865 ATCATCTATGTGCCAAACCAACAGATACAGACACGGGTGCTGGCTGTATTGTTCAAG 921

RESULT 4
US-09-685-166A-526
; Sequence 526, Application US/09685166A
; Patent No. 6630305
; GENERAL INFORMATION:
; APPLICANT: Xu, Jiangchun
; APPLICANT: Dillon, Davin C.
; APPLICANT: Mitcham, Jennifer L.
; APPLICANT: Harlocker, Susan L.
; APPLICANT: Jiang, Yuqi
; APPLICANT: Henderson, Robert A.
; APPLICANT: Kalos, Michael D.
; APPLICANT: Fanger, Gary R.
; APPLICANT: Retter, Marc W.
; APPLICANT: Stolk, John A.
; APPLICANT: Day, Craig H.
; APPLICANT: Vedvick, Thomas S.
; APPLICANT: Carter, Darrick
; APPLICANT: Li, Samuel
; APPLICANT: Wang, Aijun


```
Db 730 CTCCTTTTTCCTCCAGCTCTGCTGATCAAGCGGCTGGCTTCTGCCACTCCAAATGTC 789
Qy 184 IleLeuHisThrTyrCysGluHisMetAlaValVallysLeuAlaCysGlyAspThrArg 203
   ::::::::::::::::::::
Db 790 CTCCTGCACTCTATTGTGTCCACGAGATGAATGAAGTTGGCTATGCGACACACTTTG 849
Qy 204 ProAsnArgValTyrGlyLeuThrAlaAlaLeuLeuValIleGlyValAspLeuPheCys 223
   ::::::::::::::::::::
Db 850 CCCAATGGGTATATGGTCTTACTGCCATTCGTCTGGTCATGGCGTGAGCGTAAATGTTTC 909
Qy 224 IleGlyLeuSerTyrAlaLeuIleAlaGlnAlaValLeuArgLeuSerSerHisGluAla 243
   ::::::::::::::::::::
Db 910 ATCTCTGTCTCTATTCTGTAATACGACGGTCTCTGCAACTGCCCTTCCAAAGTCAGAG 969
Qy 244 ArgSerLysAlaLeuGlyThrCysGlySerHisValCysValIleLeuLeuSerTyrThr 263
   ::::::::::::::::::::
Db 970 CGGGCCAAAGGCTTTGGAACCTGTGTGCACACATTTGGTGTGGTACTCGCCTTCTATGTG 1029
Qy 264 ProAlaLeuPheSerPheThrHisArgPheGlyHisValProValHisIleHis 283
   ::::::::::::::::::::
Db 1030 CCACTTATGGCTCTCAGTTGTACACCGCTTTGGAACAGCTTCATCCCATTTGGCGT 1089
Qy 284 IleLeuLeuAlaAsnValTyrLeuLeuLeuProProAlaLeuAsnProValValTyrGly 303
   ::::::::::::::::::::
Db 1090 GTTGTCAATGGGTGACATCTACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1149
Qy 304 VallysThrLysGlnIleArgLysArgValValArgValPheGln 318
   ::::::::::::::::::::
Db 1150 GCCAAACCAACAGATCAGAACACCGGTGCTGGCTATGTTCAAG 1194
```

RESULT 7

```
US-09-339-115-1
; Sequence 1, Application US/09339115
; Patent No. 6372891
; GENERAL INFORMATION:
; APPLICANT: Soppet, Daniel R.
; APPLICANT: Li, Yi
; APPLICANT: Rosen, Craig A.
; APPLICANT: Ruben, Steven M.
; TITLE OF INVENTION: HUMAN G-PROTEIN RECEPTOR HPRAJ70
; NUMBER OF SEQUENCES: 8
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: CARELLA, BYRNE, BAIN, GILFILLAN, CECCHI,
; ADDRESS: STUART & OLSTEIN
; STREET: 6 Becker Farm Road
; CITY: Roseland
; STATE: New Jersey
; COUNTRY: USA
; ZIP: 07068
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/339,115
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 09/053,303
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Ferraro, Gregory D.
; REGISTRATION NUMBER: 36,134
; REFERENCE/DOCKET NUMBER: 325800-446
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 201-994-1700
; TELEFAX: 201-994-1744
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1474 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
```

```
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 274..1233
; US-09-339-115-1

Alignment Scores:
Pred. No.: 2,24e-91 Length: 1474
Score: 856.00 Matches: 160
Percent Similarity: 70.51% Conservative: 48
Best Local Similarity: 54.24% Mismatches: 87
Query Match: 49.83% Indels: 0
DB: 4 Gaps: 0

US-10-081-775-2 (1-329) x US-09-339-115-1 (1-1474)
Qy 24 LeuLeuGlyIleProGlyLeuGluGlnPheHisLeuTrpLeuSerLeuProValCysGly 43
Db 310 CTTATTGGTATCCAGGATTAGAGAAAGCCCATTTCTGGGTGGCTTCCCTCTCTTCC 369
Qy 44 LeuGlyThrAlaThrIleValGlyAsnIleThrIleLeuValValAlaThrGluPro 63
Db 370 ATGTATGTAGTGGCAATGTGTGGAACTGTGATCGTGTCTTCATCGTAAAGACGGAACGC 429
Qy 64 ValLeuHisLysProValTyrLeuPheLeuCysMetLeuSerThrIleAspLeuAlaAla 83
Db 430 AGCTGTACAGCTCGATGTACTCTTCTCTGTCATGCTTGCAGCCATTCACCTGGCCTTA 489
Qy 84 SerValSerThrValProLysLeuAlaIlePheTrpCysGlyAlaGlyHisIleSer 103
Db 490 TCCACATCCACCATGCTTAAGATCTTGCCTTTTCTGTTTGTATCCGAGAGATTAGC 549
Qy 104 AlaSerAlaCysLeuAlaHisMetPheIleHisAlaPheCysMetMetGluSerThr 123
Db 550 ATTGAGGCTGTCTTACCAGATGTCTTATTATGCGCCTCTCAGCCATTCAGCCATTCAC 609
Qy 124 ValLeuLeuAlaMetAlaPheAspArgTyrValAlaIleCysHisProLeuArgTyrAla 143
Db 610 ATCTGTCTGGCCATGGCTTTGACCGTTATGTGGCCATCTGCCACCCACTGCGCATGCT 669
Qy 144 ThrIleLeuThrAspThrIleAlaHisIleGlyValAlaAlaValAlaValArgGlySer 163
Db 670 GCAGTGTCAACAATACAGTAACAGCCAGATTTGGCATCTGGTGTGTGTCGCGGATCC 729
Qy 164 LeuLeuMetLeuProCysProPheIleGlyArgLeuAsnPheCysGlnSerHisVal 183
Db 730 CTCCTTTTTCCTCCTGCTGATCAAGCGGCTGGCCTTCTGCCACTCCAAATGTC 789
Qy 184 IleLeuHisThrTyrCysGluHisMetAlaValVallysLeuAlaCysGlyAspThrArg 203
Db 790 CTCCTGCACTCTATTGTGTCCACGAGATGAATGAAGTTGGCTATGCGACACTTTG 849
Qy 204 ProAsnArgValTyrGlyLeuThrAlaAlaLeuLeuValIleGlyValAspLeuPheCys 223
Db 850 CCCAATGGGTATATGGTCTTACTGCCATTCGTCTGGTGTGGTGGCGGTGAGCGTAAATGTTTC 909
Qy 224 IleGlyLeuSerTyrAlaLeuIleAlaGlnAlaValLeuArgLeuSerSerHisGluAla 243
Db 910 ATCTCTGTCTCTATTCTGTAATACGACGGTCTGCAACTGCCCTTCCAAAGTCAGAG 969
Qy 244 ArgSerLysAlaLeuGlyThrCysGlySerHisValCysValIleLeuLeuSerTyrThr 263
Db 970 CGGGCCAAAGGCTTTGGAACCTGTGTGCACACATTTGGTGTGGTACTCGCCTTCTATGTG 1029
Qy 264 ProAlaLeuPheSerPheThrHisArgPheGlyHisValProValHisIleHis 283
Db 1030 CCACTTATGGCTCTCAGTTGTACACCGCTTTGGAACAGCTTCATCCCATTTGGCGT 1089
Qy 284 IleLeuLeuAlaAsnValTyrLeuLeuLeuProProAlaLeuAsnProValValTyrGly 303
Db 1090 GTTGTCAATGGGTGACATCTACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1149
```

P

QY 268 SerPhePheThrHisArgPheGlyHisHisValProValHisIleHisIleLeuLeuAla 287
Db 790 ACCTATTGGAGGCCCAAGTCTAGCCAC-----TCACGAGGAATGGACAAATCTTGGCC 843
QY 288 AsnValTyrLeuLeuLeuProProAlaLeuAsnProValValTyrGlyValValValThrLys 307
Db 844 CTCTTCTACACAGTAGTGACATCATCTGCTGACCCCTATCATCTATATAGTTTAAGGAACAAG 903
QY 308 GlnIleArgLysArgValValArgValPheGlnSerGlyGlnGlyMetGlyIleLys 326
Db 904 GAAGTCAAGGCAGCAGCTGAGAGAACT-----CTGGGCCTGAAA 942
RESULT 10
US-08-748-506-7
; Sequence 7, Application US/08748506
; Patent No. 6159707
; GENERAL INFORMATION:
; APPLICANT: Ronnett et al.
; TITLE OF INVENTION: NOVEL SPERM RECEPTORS
; NUMBER OF SEQUENCES: 31
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Leydig, Voit & Mayer, Ltd.
; STREET: Two Prudential Plaza, Suite 4900
; CITY: Chicago
; STATE: IL
; COUNTRY: US
; ZIP: 60601-6780
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/748,506
; FILING DATE: 08-NOV-1996
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 60/033,751
; FILING DATE: 09-NOV-1995
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION: 74940
; REFERENCE/DOCKET NUMBER:
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 312-616-5600
; TELEFAX: 312-616-5700
; INFORMATION FOR SEQ ID NO: 7:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 966 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
US-08-748-506-7
Alignment Scores:
Pred. No.: 1,81e-40 Length: 966
Score: 423.50 Matches: 105
Percent Similarity: 52.00% Conservative: 51
Best Local Similarity: 35.00% Mismatches: 133
Query Match: 24.65% Indels: 11
DB: 3 Gaps: 4
US-10-081-775-2 (1-329) x US-08-748-506-7 (1-966)
QY 27 IleProGlyLeuGluGlnPheHisLeuTrpLeuSerLeuProValCysGlyLeuGlyThr 46
Db 76 GTCCCTGGA--GAATGCTTCTCTGTTACCCCTCACTCTCTCATGTTCTTA-----126
QY 47 AlaThrIleValGlyAsnIleThrIleLeuValValAlaThrGluProValLeuHis 66
Db 127 GTATCACTAACAGGAATACTCTCATAGCCCTTGCTATTGTACCACTGCTCTCTACAC 186
QY 67 LysProValTyrLeuPheLeuCysMetLeuSerThrIleAspLeuAlaIleValSer 86

CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 60/033,751
FILING DATE: 09-NOV-1995
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION: 74940
REFERENCE/DOCKET NUMBER:
TELECOMMUNICATION INFORMATION:
TELEPHONE: 312-616-5600
TELEFAX: 312-616-5700
INFORMATION FOR SEQ ID NO: 5:
SEQUENCE CHARACTERISTICS:
LENGTH: 966 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
US-08-748-506-5
Alignment Scores:
Pred. No.: 1.58e-40 Length: 966
Score: 424.00 Matches: 96
Percent Similarity: 51.61% Conservative: 48
Best Local Similarity: 34.41% Mismatches: 127
Query Match: 24.68% Indels: 8
DB: 3 Gaps: 2
US-10-081-775-2 (1-329) x US-08-748-506-5 (1-966)
QY 48 ThrIleValGlyAsnIleThrIleLeuValValAlaThrGluProValLeuHisLys 67
Db 130 TCATTAACAGGAATACTCTCATGATCTCTCTCTGGAGATTGGCTATCTATCTCTCTCTC 189
QY 68 ProValTyrLeuPheLeuLeuMetLeuSerThrIleAspLeuAlaIleValSerThr 87
Db 190 CCCATGTAATCTTTCTGCGCAACTTGTCTCTCTGGAGATTGGCTATCTATCTCTCTCTC 249
QY 88 ValProLysLeuLeuAlaIlePheTrpCysGlyAlaGlyHisIleSerAlaSerAlaCys 107
Db 250 ATACCAAGATGCTGTCAGAGCCTTGTGAGTGAGGCCAGAGATCTCTCGGAGGAGTGT 309
QY 108 LeuAlaHisMetPheIleHisAlaPheCysMetMetGluSerThrValLeuLeuAla 127
Db 310 GCCACACAGATGTTTTTTTTCGATTTTGGTATTAATGATGCTGCTTATGGCAGCC 369
QY 128 MetAlaPheAspArgTyrValAlaIleCysHisProLeuArgTyrAlaThrIleLeuThr 147
Db 370 ATGGCCTTTCACCGCTGCGATGCCATATGCTCCCACTCCCATATGCAACCGAATGAGT 429
QY 148 AspThrIleIleAlaHisIleGlyValAlaAlaValValArgGlySerLeuLeuMetLeu 167
Db 430 CCGAGAGTATGTCGCCATTTGGCAATTTGTTTCATGGGGAATGGGATGATAGTAGTCTG 489
QY 168 ProCysProPhePheIleGlyArgLeuAsnPheCysGlnSerHisValIleLeuHisThr 187
Db 490 GGACAAACCAATTTATTTCTCTTGAATCTCTGTGACCCCTGTGAAATAGACCACTTC 549
QY 188 TyrCysGluHisMetAlaValLysLeuAlaCysGlyAspThrArgProAsnArgVal 207
Db 550 TTCTGTGACCTTCCACCTCTCTGCGCACTTGTGCTGTGGAGATACATCCCAAAACGAGGCT 609
QY 208 TyrGlyLeuThrAlaAlaLeuLeuValIleGlyValAspPheCysIleGlyLeuSer 227
Db 610 GCCATCTTTGGTGGATGCTCTGCTATCTATCTAGCCCAATTTTGTGATCATTTATTTCT 669
QY 228 TyrAlaLeuIleAlaGlnAlaValLeuArgLeuSerSerHisGluAlaArgSerLysAla 247
Db 670 TATGTCAAATTTCTCATTCAGTGTCTTCTGATGCTTACCTTGAGGGGCGCCATAAGCT 729
QY 248 LeuGlyThrCysGlySerHisValCysValIleLeuIleSerTyrThrProAlaLeuPhe 267
Db 730 CTTTCCACCTGTGCTCTACCTTACTTGTAGTCACACTTTTATGCTCAGCATGTATT 789

```
Db 187 ACCCCCATGACTCTCTTCTGGCCAACTTGTCTCTCTGGAGATTGGCTATCTGCTCT 246
Qy 87 ThrValProLysLeuLeuAlaIlePheTrpCysGlyAlaGlyHisIleSerAlaSerAla 106
Db 247 GTCATACCCAGATCTCTCTGAGCCCTGTGAGTGGAGCCGAGGATCTCTGGAGGGT 306
Qy 107 CysLeuAlaHisMetPhePheIleHisAlaPheCysMetMetGluSerThrValLeuLeu 126
Db 307 TGTGCTCACAGATCTCTCTCTATATCTTGTGTATTAAGTGTGCTGCTATTTGGCA 366
Qy 127 AlaMetAlaPheAspArgTyrValAlaIleCysHisProLeuArgTyrAlaThrIleLeu 146
Db 367 GCCATGGCTTGTACCGCTATATGCTATATGCTATGCTATGCTATGCTATGCTATGCT 426
Qy 147 ThrAspThrIleIleAlaHisIleGlyValAlaValAlaValArgGlySerLeuLeuMet 166
Db 427 AGTCGTGGGTATGCTGCTATTTGGCAATGCTCTGCTGCTGCTGCTGCTGCTGCTGCT 486
Qy 167 LeuProCysProPhePheIleGlyArgLeuAsnPheCysGlnSerHisValIleLeuHis 186
Db 487 CTGGACAGACCAATTTTATTTCTCTTGAATCTCTGCTGCTGCTGCTGCTGCTGCTGCT 546
Qy 187 ThrTyrCysGluHisMetAlaValValLysLeuAlaCysGlyAspThrArgProAsnArg 206
Db 547 TTCTTCTGTGACCTTCCACCTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 606
Qy 207 ValTyrGlyLeuThrAlaAlaLeuValIleGlyValAlaLeuAspLeuPheCysIleGlyLeu 226
Db 607 GCTGCCATCTTTGGGAGCAGTGTCTGCAATTTAGTCCATTTTACTGATCATTTCT 666
Qy 227 SerTyrAlaLeuIleAlaGlnAlaValLeuArgLeuSerHisGluAlaArgSerLys 246
Db 667 TCCTATGTCAGAAATCTCGTGTGAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 726
Qy 247 AlaLeuGlyThrCysGlySerHisValCysValIleLeuLeuSerTyrThrProAlaLeu 266
Db 727 GCTCTCTCTACCTGCTCATCTCACTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 786
Qy 267 PheSerPhePheThrHisArgPheGlyHisValProValHisIleHisIleLeuLeu 286
Db 787 GCCACCTTTGAGTCAAGTCTAGCCATCTAGCCATCTAGCCATCTAGCCATCTAGCCATCT 840
Qy 287 AlaAsnValTyrLeuLeuProAlaLeuAsnProValValTyrValTyrGlyValLysThr 306
Db 841 GCCCTCTCTATACATCAGTGCATCCATGCTGATCCATCCATCCATCCATCCATCCATCC 900
Qy 307 LysGlnIleArgLysArgValValArgValPheGlnSerGlyGlnGlyMetGlyIleLys 326
Db 901 AAGGAAGTAAGGGTGTCACTGAGAGAACT-----CTGGGCTGAAG 942
```

RESULT 11

```
US-08-748-506-6
; Sequence 6, Application US/08748506
; Patent No. 6159707
; GENERAL INFORMATION:
; APPLICANT: Ronnett et al.
; TITLE OF INVENTION: NOVEL SPERM RECEPTORS
; NUMBER OF SEQUENCES: 31
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Leydig, Voit & Mayer, Ltd.
; STREET: Two Prudential Plaza, Suite 4900
; CITY: Chicago
; STATE: IL
; COUNTRY: US
; ZIP: 60601-6780
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: IBM PC compatible
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/748,506
```

```
; FILING DATE: 08-NOV-1996
; CLASSIFICATION: 435
; PRIOR APPLICATION NUMBER: US 60/033,751
; FILING DATE: 09-NOV-1995
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; REFERENCE/DOCKET NUMBER: 74940
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 312-616-5600
; TELEFAX: 312-616-5700
; INFORMATION FOR SEQ ID NO: 6:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 966 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
US-08-748-506-6

Alignment Scores:
Pred. No.: 1,22e-39 Length: 966
Score: 416.50 Matches: 102
Percent Similarity: 50.17% Conservative: 48
Best Local Similarity: 34.11% Mismatches: 144
Query Match: 24.24% Indels: 5
DB: 3 Gaps: 3

US-10-081-775-2 (1-329) x US-08-748-506-6 (1-966)
```

```
Qy 27 IleProGlyLeuGluInPheHisLeuTrpLeuSerLeuProValCysGlyLeuGlyThr 46
Db 76 GTCCTTGA--GAATGCTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 126
Qy 47 AlaThrIleValGlyAsnIleThrIleLeuValValValAlaThrGluProValLeuHis 66
Db 127 GTATCACTACAGAAATGCTCTCATAGCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 186
Qy 67 LysProValTyrLeuPheLeuCysMetLeuSerThrIleAspLeuAlaAlaSerValSer 86
Db 187 ACCCCCATGACTCTCTTCTGGCCAACTTGTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 246
Qy 87 ThrValProLysLeuLeuAlaIlePheTrpCysGlyAlaGlyHisIleSerAlaSerAla 106
Db 247 GTCATACCCAGATCTCTCTGAGCCCTGTGAGTGGAGCCGAGGATCTCTGGAGGGT 306
Qy 107 CysLeuAlaHisMetPhePheIleHisAlaPheCysMetMetGluSerThrValLeuLeu 126
Db 307 TGTGCCACAGATCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 366
Qy 127 AlaMetAlaPheAspArgTyrValAlaIleCysHisProLeuArgTyrAlaThrIleLeu 146
Db 367 GCCATGGCTTGTACCGCTATATGCTATATGCTATGCTATGCTATGCTATGCTATGCTATG 426
Qy 147 ThrAspThrIleIleAlaHisIleGlyValAlaValAlaValArgGlySerLeuLeuMet 166
Db 427 AGTCGTGAGTATGTGCCAATTTGGCAATTTAGTCCATTTTACTGATCATTTCTGCTAT 486
Qy 167 LeuProCysProPhePheIleGlyArgLeuAsnPheCysGlnSerHisValIleLeuHis 186
Db 487 CTGGACAGACCAATTTTATTTCTCTTGAATCTCTGCTGCTGCTGCTGCTGCTGCTGCT 546
Qy 187 ThrTyrCysGluHisMetAlaValValLysLeuAlaCysGlyAspThrArgProAsnArg 206
Db 547 TTCTTCTGTGACCTTCCACCTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 606
Qy 207 ValTyrGlyLeuThrAlaAlaLeuValIleGlyValAlaLeuPheCysIleGlyLeu 226
Db 607 GCTGCCATCTTTGGGAGCAGCAATCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 666
Qy 227 SerTyrAlaLeuIleAlaGlnAlaValLeuArgLeuSerHisGluAlaArgSerLys 246
Db 667 TCTTATGTGAGAAATCTCGTGTGAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 726
```

QY	247	AlaLeuGlyThrCysGlySerHisValCysValIleLeuLeuSerTyrThrProAlaLeu	266
Db	727	GCTCTTCCACCTGCTCTCCTCAGACCTACTTGTAGTCACACTCTTTTATGGCTCTGTCTCC	786
QY	267	PheSerPhePheThrHisArgPheGlyHisValProValHisIleHisIleLeuLeu	286
Db	787	TTTACCTATTGAGGCCAAGTCTAGCCAC-----TCACCAGAAATGGACAACTCTTG	840
QY	287	AlaAsnValTyrLeuLeuLeuProAlaLeuAsnProValValTyrGlyValTyrThr	306
Db	841	GCCCTCTTACACAGCAGTGCATCCATGCTGAACCTCATCTACAGCTCTAAGGAAC	900
QY	307	LysGlnIleArgLysArgValValArgValPheGlnSerGlyGlnGlyMetGlyLe	325
Db	901	AAGGAAGTCAAGGCAGCAGTCTGAGGAAGTCTCGACCTGAAAAAATATATGCAATT	957
RESULT 12			
US-08-748-506-8			
; Sequence 8, Application US/08748506			
; Patent No. 6159707			
; GENERAL INFORMATION:			
; APPLICANT: Ronnett et al.			
; TITLE OF INVENTION: NOVEL SPERM RECEPTORS			
; NUMBER OF SEQUENCES: 31			
; CORRESPONDENCE ADDRESS:			
; ADDRESSEE: Leydig, Voit & Mayer, Ltd.			
; STREET: Two Prudential Plaza, Suite 4900			
; CITY: Chicago			
; STATE: IL			
; COUNTRY: US			
; ZIP: 60601-6780			
; COMPUTER READABLE FORM:			
; MEDIUM TYPE: Floppy disk			
; COMPUTER: IBM PC compatible			
; OPERATING SYSTEM: PC-DOS/MS-DOS			
; SOFTWARE: PatentIn Release #1.0, Version #1.25			
; CURRENT APPLICATION DATA:			
; FILING DATE: 08-NOV-1996			
; APPLICATION NUMBER: US/08/748,506			
; CLASSIFICATION: 435			
; PRIOR APPLICATION DATA:			
; APPLICATION NUMBER: US 60/033,751			
; FILING DATE: 09-NOV-1995			
; CLASSIFICATION: 435			
; ATTORNEY/AGENT INFORMATION:			
; REFERENCE/DOCKET NUMBER: 74940			
; TELECOMMUNICATION INFORMATION:			
; TELEPHONE: 312-616-5600			
; TELEFAX: 312-616-5700			
; INFORMATION FOR SEQ ID NO: 8:			
; SEQUENCE CHARACTERISTICS:			
; LENGTH: 966 base pairs			
; TYPE: nucleic acid			
; STRANDEDNESS: double			
; TOPOLOGY: linear			
; MOLECULE TYPE: DNA (genomic)			
US-08-748-506-8			
Alignment Scores:			
Pred. No.:	6,27e-39	Length:	966
Score:	410.50	Matches:	99
Percent Similarity:	50.6%	Conservative:	48
Best Local Similarity:	34.14%	Mismatches:	138
Query Match:	23.89%	Indels:	5
DB:	3	Gaps:	3
US-10-081-775-2 (1-329) x US-08-748-506-8 (1-966)			
QY	27	IleProGlyLeuGlnPheHisIleTyrLeuSerLeuProValCysGlyLeuGlyThr	46
Db	76	GTCCCTGGA---GAATGCTTCTCTCTGTTTCAACCTCATCTCTATGTTTAAAGAAC	126

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Word Perfect 6.1 for Windows/MS-DOS 6.2
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/016,434
FILING DATE: HERewith

CLASSIFICATION:
PRIORITY APPLICATION DATA:
APPLICATION NUMBER:
FILING DATE:
CLASSIFICATION:
ATTORNEY/AGENT INFORMATION:
NAME: Zeller, Karen J.
REGISTRATION NUMBER: 37,071
REFERENCE/DOCKET NUMBER: PA-0002 US
TELECOMMUNICATION INFORMATION:
TELEPHONE: (650) 855-0555
TELEFAX: (650) 845-4166
INFORMATION FOR SEQ ID NO: 1413:
SEQUENCE CHARACTERISTICS:
LENGTH: 1282 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
IMMEDIATE SOURCE:
LIBRARY: GENBANK
CLONE: G516319
US-09-016-434-1413

Alignment Scores:
Pred. No.: 1,12e-38 Length: 1282
Score: 410.00 Matches: 97
Percent Similarity: 51.00% Conservative: 56
Best Local Similarity: 32.33% Mismatches: 135
Query Match: 23.86% Indels: 12
DB: 4 Gaps: 4

US-10-081-775-2 (1-329) x US-09-016-434-1413 (1-1282)

QY 22 PhePheLeuLeuGlyLeu-----ProGlyLeuGluGln-----PheHisLeuTirp 36
Db 201 TTATCTCTGCTGGCTTGTGGAGGCGCCAGGCTGCAGCCAGTGTCTTGTCTTCTTC 260
QY 37 LeuSerLeuProValCysGlyLeuGlyThraAlaThrileValGlyAsnileThrileLeu 56
Db 261 CTCTTTGCCTTACCTG-----GTCACGCTCAGGGGCAACCTCAGCATCTG 305
QY 57 ValValValAlaThrGluProValLeuHisLysProValTyrLeuPheLeuCysMetLeu 76
Db 306 GCAGCTGCTTGTGGAGCCCAACTCCACCCCACTGCTTCTTCTGGGAACTA 365
QY 77 SerThrileAspLeuAlaSerValSerThrValProLysLeuLeuAlailePheTirp 96
Db 366 TCAGTGTGCTGATGTGGGTGCATCAGCGTCACTGCTTCCATCAATGTGAGTCTCTCTG 425
QY 97 CysGlyAlaGlyHisIleSerAlaSerAlaCysLeuAlaHisMetPhePheIleHisAla 116
Db 426 TCCCGCAAGCTGAGTTCCTGTGGGGCTGCTTACCCAGCTCTTCTTCTTCCATCTG 485
QY 117 PheCysMetMetGluSerThrValLeuAlaMetAlaPheAspArgTyrValAlaile 136
Db 486 TTCTGTGAGTGGAGTCTTCTGCTGACCGCATGCGCTATGACCAATTCCTGGGCATC 545
QY 137 CysHisProLeuArgTyrAlaThrileLeuLeuThrileLeuAlaHisIleGlyVal 156
Db 546 TGGCGGCCCCCTACCTACAGCCCGCATGAGTGCAGACATCCAGAGGATGTGGTGGCT 605
QY 157 AlaAlaValValargGlySerLeuLeuMetLeuLeuProCysProPhePheIleGlyA-gLeu 176
Db 606 GCCTCTCTGGGCTGTGCTTTTACCAACAGCACTGACCCACACACTGCGCCATGTCCAGCTC 665

QY 177 AsnPheCysGlnSerHisValIleLeuHisThrTyrCysGluHisMetAlaValVallys 196
Db 666 AACTTCTGTGGCCCAATGTGATCAATCACTTCTACTGTGACCTCCACAGCTCTTCCAG 725
QY 197 LeuAlaCysGlyAspThrArgProAsnArgValTyrGlyLeuThrAlaAlaLeuVal 216
Db 726 CTCTCTGCTGCCAGCACCACTCAATGAGCTGTGCTCTTTTGTGTGGTGTATAATG 785
QY 217 IleGlyValAspLeuPheCysIleGlyLeuSerTyrAlaLeuIleAlaGlnAlaValLeu 236
Db 786 GCAGGTACCCCATGCTCTCATTTCTATCTATATCCAGTGCAGCTGCAGTCCCTG 845
QY 237 ArgLeuSerSerHisGluAlaArgSerLysAlaLeuGlyThrCysGlySerHisValCys 256
Db 846 CGAATTCGCTCTGTAGAGGGCAGGAAGAAAGCTTCTCCACATGTGGTCCACCTCACT 905
QY 257 ValIleLeuIleSerTyrThrProAlaLeuPheSerPhePheThrHisArgPheGlyHis 276
Db 906 GTGGTGGCATATCTTATGTTGAGTATCTTAACTATATG-----CGACTGGTTC 959
QY 277 HisValProValHisIleHisIleLeuLeuAlaAsnValTyrLeuLeuProAla 296
Db 960 ACCAAGCTTTTCAGACAAGGATAAAGCTGTTGGAATTTTCAACACTGTCTCAATCCCATG 1019
QY 297 LeuAsnProValValTyrGlyValLysThrLysGlnIleArgLysArgValValArgVal 316
Db 1020 CTGAACCAATCATCTACAGCTTCAGAAACCTGATGTGAGAGTGCCATCTGGAGGATG 1079

RESULT 14
US-09-546-986A-5
; Sequence 5, Application US/09546986A
; Patent No. 6635741
; GENERAL INFORMATION:
; APPLICANT: Powers, Scott
; APPLICANT: Yang, Jianxin
; APPLICANT: Cutler, Gene
; APPLICANT: Tularik Inc.
; TITLE OF INVENTION: No. 6635741el G-Protein Coupled Receptors
; FILE REFERENCE: 018781-004720US
; CURRENT APPLICATION NUMBER: US/09/546,986A
; CURRENT FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 09/524,730
; PRIOR FILING DATE: 2000-03-14
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 5
; LENGTH: 1351
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (86)..(1108)
; OTHER INFORMATION: human breast cancer amplified G-protein coupled
; OTHER INFORMATION: receptor 3 (BCA-GPCR-3)
US-09-546-986A-5

Alignment Scores:
Pred. No.: 1,21e-38 Length: 1351
Score: 410.00 Matches: 101
Percent Similarity: 48.62% Conservative: 57
Best Local Similarity: 31.08% Mismatches: 131
Query Match: 23.86% Indels: 36
DB: 4 Gaps: 5

US-10-081-775-2 (1-329) x US-09-546-986A-5 (1-1351)

QY 6 GlyHisAsnMetGluSerProHisHisThrAspValAsp-----Pro 19
Db 116 GTGGGCTTTTGGCCCCCAGCAGCATACATGATGGAATAGCAATGTGAGTCTTCCA 175
QY 20 SerValPhePheLeuGlyIleProGlyLeuGluGlnPheHisLeuTrpLeuSerLeu 39
Db 176 GAAGTCTTGTCTCTGCTGGGCTTCTCCGACGACCCCTACTAGAAACTGTCTCTTCTATA 235


```

Db      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
812 GCCCGCGCGTGTGAAGATCAGGTACAGCAGAGCGCGAGAGGCATTCAACACCTGT 871
QY      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
252 GlySerHisValCysValIleLeuIleSerTyrThrProAlaLeuPheSerPhePhe--- 270
Db      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
872 TCTTCCACGCTGGCTGTGTCTCTCTTTTACGGGAGCATCATCTTCATGTATCTCCAG 931
QY      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
271 -----ThrHisArgPheGlyHisHisValProValHisIleLeu 285
Db      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
932 CCAGCCAGAGCACCTCCCATGAGCAGGCAAG-----TTC 967
QY      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
286 LeuAlaAsnValTyrLeuLeuLeuProProAlaLeuAsnProValValTyrGlyVallys 305
Db      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
968 ATAGCTCTGTCTTACACCGTAGTCACTCTCGGTGAGCCCACTTATTACACCCCTGAGG 1027
QY      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
306 ThrLysGlnIleArg 310
Db      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
1028 AACCGAGGTGAAG 1042

```

Search completed: August 27, 2004, 20:57:49
 Job time : 98 secs

This Page Blank (uspo)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: August 27, 2004, 18:28:44 ; Search time 3045 Seconds
(without alignments)
3226.488 Million cell updates/sec

Title: US-10-081-775-2
Perfect score: 1718
Sequence: 1 MSSTLGHNMSPHPTDVPD.....RKRVRVRFSGQGMGIKASE 329

Scoring table: BLOSUM62

Xgapop 10.0, Xgapext 0.5
Ygapop 10.0, Ygapext 0.5
Fgapop 6.0, Fgapext 7.0
Delop 6.0, Delext 7.0

Searched: 27513289 seqs, 14931090276 residues

Total number of hits satisfying chosen parameters: 55026578

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-O=/cgn2_1/USPTO_spool/US10081775/runat_20082004_170213_11139/app_query.fasta_1.519
-DB=FAST_QFMT=fastap -SUFFIX=first -MINMATCH=0.1 -LOOPEXT=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45
-DOCALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US10081775 @CGN 1 1 3437 @runat_20082004_170213_11139 -NCPUP=6 -ICPU=3
-NO.MMAP -LARGQUERY -NEG SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOC
-DEV TIMEOUT=120 -WARN TIMEOUT=40 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

1: em_estba.*
2: em_esthum.*
3: em_estin.*
4: em_estmu.*
5: em_estov.*
6: em_estpl.*
7: em_estro.*
8: em_hcc.*
9: gb_est1.*
10: gb_est2.*
11: gb_hcc.*
12: gb_est3.*
13: gb_est4.*
14: gb_est5.*
15: em_estfun.*
16: em_estom.*
17: em_gss_hum.*
18: em_gss_inv.*
19: em_gss_pln.*
20: em_gss_vrt.*
21: em_gss_fun.*
22: em_gss_man.*
23: em_gss_mus.*
24: em_gss_pro.*
25: em_gss_rod.*
26: em_gss_phg.*
27: em_gss_vri.*
28: gb_gss1.*

29: gb_gss2.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB	ID	Description
1	977.5	56.9	948	29	AY401474	Mus muscu
2	961.5	56.0	948	29	AY401472	Mus muscu
3	904	52.6	936	29	AY401478	Mus muscu
4	889.5	51.8	948	29	AY401473	Mus muscu
5	889.5	51.8	957	29	AY414225	Mus muscu
6	882.5	51.4	954	29	AY414227	Mus muscu
7	880.5	51.3	936	29	AY401480	Mus muscu
8	868	50.5	1964	11	AK036356	Mus muscu
9	868	50.5	3410	11	AK028467	Mus muscu
10	867.5	50.5	948	29	AY410600	Mus muscu
11	834	48.5	948	29	AY410598	Mus muscu
12	828.5	48.2	846	29	AY402272	Mus muscu
13	827	48.1	948	29	AY410599	Mus muscu
14	819	47.7	966	29	AY405569	Mus muscu
15	802	46.7	1086	28	AF101706	Mus muscu
16	794	46.2	822	29	CC500683	Mus muscu
17	788	45.9	966	29	AY405567	Mus muscu
18	784	45.6	900	29	AY410597	Mus muscu
19	769	44.8	900	29	AY410595	Mus muscu
20	768.5	44.7	846	29	AY402270	Mus muscu
21	762	44.4	900	29	AY410596	Mus muscu
22	750	43.7	774	29	CC512357	Mus muscu
23	730.5	42.5	746	29	AY401479	Mus muscu
24	730	42.5	948	29	AY402752	Mus muscu
25	717.5	41.8	957	29	AY414226	Mus muscu
26	717.5	41.5	939	29	AY402268	Mus muscu
27	704.5	41.0	798	28	BH082934	Mus muscu
28	703.5	40.9	939	29	AY402267	Mus muscu
29	696	40.5	963	29	AY401651	Mus muscu
30	678.5	39.5	853	29	CC527411	Mus muscu
31	673.5	39.2	966	29	AY401649	Mus muscu
32	671	39.1	864	29	CC480759	Mus muscu
33	669.5	39.0	820	28	BH046083	Mus muscu
34	663.5	38.6	987	29	AY402269	Mus muscu
35	649.5	37.8	825	28	AF156720	Mus muscu
36	640	37.3	942	29	AY402281	Mus muscu
37	630	36.7	782	29	CC534047	Mus muscu
38	609.5	35.5	795	28	BZ256668	Mus muscu
39	604.5	35.2	859	28	BZ221157	Mus muscu
40	602.5	35.1	927	29	AY401477	Mus muscu
41	602	35.0	758	29	CE341239	Mus muscu
42	601.5	35.0	669	29	CC569584	Mus muscu
43	601.5	35.0	936	29	AY402284	Mus muscu
44	595	34.6	939	29	AY402279	Mus muscu
45	592	34.5	786	29	AY414230	Mus muscu

ALIGNMENTS

RESULT 1
AY401474
LOCUS
DEFINITION
AY401474
genomic survey sequence.
ACCSSION
AY401474.1 GI:39757463
VERSION
AY401474.1
KEYWORDS
GSS.
SOURCE
Mus musculus (house mouse)
ORGANISM
Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE
1 (bases 1 to 948)
AY401474
Mus musculus HCM0902 gene, VIRTUAL TRANSCRIPT, partial sequence,
948 bp
DNA
linear
GSS 12-DEC-2003

QY	189	CysGluHisMetAlaValVallyLeuAlaCysGlyAspThrArgProAsnArgValTyr	208
DB	538	TGTAGCACATGGCTGTGTGGTGAAGCTGGTGTGGGGGACACACAAACAAACATATAT	537
QY	209	GlyLeuThrAlaAlaLeuValIleGlyValAspLeuPheCysIleGlyLeuSerTyr	228
DB	598	GGCATCAGACGACCACTGGTGGTAGGAAGTACTCCATTGTATGGCATATCCTAT	657
QY	229	AlaLeuIleAlaGlnAlaValLeuArgLeuSerSerHisGluAlaArgSerIysAlaLeu	248
DB	658	GCACCTCATCTCCGGGCTGTGTGGGCGCTCTCCTCCAAAGAGGCGGCAAGAGACCTTT	717
QY	249	GlyThrCysGlySerHisValCysValIleLeuIleSerTyrThrProAlaLeuPheSer	268
DB	718	GGCACTGTGGCTGCCACCTGGGTGTCATCCTTCTTATACACAGGCTCTCTCTCA	777
QY	269	PhePheThrHisArgPheGlyHisHisValProValHisIleHisIleLeuLeuAlaAsn	288
DB	778	TTCATACCCAGCGCTTTGGTCAACATGTACCCAGGCGCTCCACATCTCTCTGGCTGAC	833
QY	289	ValTyrIleLeuLeuProProAlaLeuAsnProValValTyrGlyValIleThrLysGln	308
DB	838	CTCTACTCTGCTGGCGGCCATGCTCAACCCCATCATCTATGGATGAAAAACACAGCAG	897
QY	309	IleArgLysArgValValArgValPheGlnSerGly	320
DB	898	ATCCGATGGAGCACTCCGACTACTCAAGAGGGGT	933
RESULT 2	AY401472	948 bp DNA linear	GSS 12-DEC-2003
LOCUS	Homo sapiens HCM0902 gene, VIRTUAL TRANSCRIPT, partial sequence,		
DEFINITION	Genomic survey sequence.		
ACCESSION	AY401472		
VERSION	AY401472.1	GI:39757461	
KEYWORDS	GSS.		
SOURCE	Homo sapiens (human)		
ORGANISM	Homo sapiens		
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
AUTHORS	Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.		
1	(bases 1 to 948)		
2	(bases 1 to 948)		
AUTHORS	Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejariwal,A., Todd,M.A., Tanenbaum,D.M., Civello,D.R., Lu,F., Murphy,B., Ferriera,S., Wang,G., Zheng,X.H., White,T.J., Sninsky,J.J., Adams,M.D. and Cargill,M.		
TITLE	Inferring nonneutral evolution from human-chimp-mouse orthologous gene trios		
JOURNAL	Science 302 (5652), 1960-1963 (2003)		
PUBMED	14671302		
REFERENCE	2 (bases 1 to 948)		
AUTHORS	Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejariwal,A., Todd,M.A., Tanenbaum,D.M., Civello,D.R., Lu,F., Murphy,B., Ferriera,S., Wang,G., Zheng,X.H., White,T.J., Sninsky,J.J., Adams,M.D. and Cargill,M.		
TITLE	Direct Submision		
JOURNAL	Submitted (16-NOV-2003) Celera Genomics, 45 West Gude Drive, Rockville, MD 20850, USA		
COMMENT	This sequence was made by sequencing genomic exons and ordering them based on alignment.		
FEATURES	Location/Qualifiers		
source	1..948		
gene	/organism="Homo sapiens" /mol_type="genomic DNA" /db_xref="taxon:9606" <1..>948 /locus_tag="HCM0902"		
ORIGIN			
Alignment Scores:			
Pred. No.:	9,66e-100	Length:	948
Score:	961.50	Matches:	178
Percent Similarity:	74.04%	Conservative:	53
Rest Local Similarity:	57.05%	Mismatches:	80

Query Match:	55.97%	Indels:	1
DB:	29	Gaps:	1
US-10-081-775-2 (1-329) x AY401472 (1-948)			
QY	9	MetGluSerProHisHisThrAspValAspProSerValPhePheLeuLeuGlyIlePro	28
DB	1	ATGAAACTCATAAACCATGACCATGACCAACCAACCTCTTCTGCTCATGGGAATTC	60
QY	29	GlyLeuGluGlnPheHisLeuTrpLeuSerLeuProValCysGlyLeuGlyThrAlaThr	48
DB	61	GGCCCGAGGACATCCACATTTGGATTGCTTTCTCTGCTTCCATGATGCTCCGGCA	120
QY	49	IleValGlyAsnIleThrIleValValAlaThrGluProValLeuHisLysPro	68
DB	121	GTGCTGGGAAACATGT	180
QY	69	ValTyrLeuPheLeuCysMetLeuSerThrIleAspLeuAlaAlaSerValSerThrVal	88
DB	181	ATGTACCTGTCTCTGATGCTATCCACATGACCTGTCTCTGACCTCCACTG	240
QY	89	ProLysLeuAlaIlePheTrpCysGlyAlaGlyHisIleSerAlaSerAlaCysLeu	108
DB	241	CCCAAGCTCTTGCACATTTTGGGCAAGGATGCTGAGATCAACTTTGGGCTGTGCT	300
QY	109	AlaHisMetPhePheIleHisAlaPheCysMetGluSerThrValLeuAlaMet	128
DB	301	GCCAGATGTTCTTATCCATGCTCTCAGCTGTAGATCTGGTATCTGCTAGCAATG	360
QY	129	AlaPheAspArgTyrValAlaIleCysHisProLeuArgTyrAlaThrIleLeuThrAsp	148
DB	361	GCCTTTGACCGCTAGTACCATTTGCTGGCTCTGCACTATGGTCTCTCCCA	420
QY	149	ThrIleIleAlaHisIleGlyValAlaValAlaValArgGlySerLeuLeuMetLeuPro	168
DB	421	GAGTCTGAGGCAAGCTGGGGCTGCACGCGGCTCTGCTGTTGGGACTCATGACCCCA	480
QY	169	CysProPhePheIleGlyArgLeuAsnPheCysGlnSerHisValIleLeuHisThrTyr	188
DB	481	CTCACCTCTTACTGGCAGACTGAGCTACTGCTCTCTCTCTCTCTCTCTCTCTCTCT	537
QY	189	CysGluHisMetAlaValLysLeuAlaCysGlyAspThrArgProAsnArgValTyr	208
DB	538	TGTGAACACATGGCTGTGTAAGCTGCTGTGGAGGAAACACAGCCCAACATCTAT	597
QY	209	GlyLeuThrAlaAlaLeuValIleGlyValAspLeuPheCysIleGlyLeuSerTyr	228
DB	598	GGCATCTGCTGCCACACTGGTGGGGCACTGACTCTCTCTCTCTCTCTCTCTCTCTCT	657
QY	229	AlaLeuIleAlaGlnAlaValLeuArgLeuSerSerHisGluAlaArgSerLysAlaLeu	248
DB	658	GCATCTATCTCCGAGCTGTGTAGGCTTTCTCCAGGAGGCAAGGCTTAAGACCTTT	717
QY	249	GlyThrCysGlySerHisValCysValIleLeuLeuSerTyrThrProAlaLeuPheSer	268
DB	718	GGCACTTGTGGCTCCCACTGGGTGTATCTCTCTCTCTCTCTCTCTCTCTCTCTCT	777
QY	269	PhePheThrHisArgPheGlyHisValProValHisIleHisIleLeuLeuAlaAsn	288
DB	778	TTCTACACACAGCGGTTTGGCCACACGTCGCCCGGCACATCCACATCTCTCTAGCTGAC	837
QY	289	ValTyrLeuLeuLeuProProAlaLeuAsnProValValTyrGlyValLysThrLysGln	308
DB	838	CTCTACCTGGTGTGGCCACCATCTCTCAACCCCATCTCTCTCTCTCTCTCTCTCTCTCT	368
QY	309	IleArgLysArgValValArgValPheGlnSerGly	320
DB	898	ATCTGGGATGGGGCTCTCGGCTCTCTGAAGTGGGGC	933
RESULT 3			
AY401478			
LOCUS			
DEFINITION Homo sapiens HCM0904 gene, VIRUAL TRANSCRIPT, partial sequence,			
936 bp DNA linear GSS 12-DEC-2003			
US-10-081-775-2 (1-329) x AY401478 (1-936)			
QY	7	HisAsnMetGluSerProHisHisThrAspValAspProSerValPhePheLeuLeuGly	26
DB	10	CACAAACAGACATATTTTAC-----CCAGTCACATTTTCTCTCATGGA	54
QY	27	IleProGlyLeuGluGlnPheHisLeuTrpLeuSerLeuProValCysGlyLeuGlyThr	46
DB	55	ATCCAGGCTCTGGAAGACTTCCACATGTGGATCTCCGGGCTTCTGCTCTGTACCTT	114
QY	47	AlaThrIleValGlyAsnIleThrIleLeuValValAlaThrGluProValLeuHis	66
DB	115	GTGCTTGTGGCAATGTCACCATCTCTGTAGTATCATCAAGGTAGAACAGACTCTCGG	174
QY	67	LysProValTyrLeuPheLeuCysMetLeuSerThrIleAspLeuAlaAlaSerValSer	86
DB	175	GAGCCCAATGTTCTCTCTGGCCATCTTTTCCACTATTCATTTGGGCTTCTCTGCAACC	234
QY	87	ThrValProLysLeuLeuAlaIlePheTrpCysGlyAlaGlyHisIleSerAlaSerAla	106
DB	235	TCGTGCTCGCATGTGGGTATCTCTGTTGATGCTCAGAGATTAATATGAGGCT	294
QY	107	CysLeuAlaHisMetPhePheIleHisAlaPheCysMetMetGluSerThrValLeuLeu	126
DB	295	TGTGTGCCCGAGATGTTTCTGATCCATGCTTCACTGGCATGGGCTGAGGCTTACTG	354
QY	127	AlaMetAlaPheAspArgTyrValAlaIleCysHisProLeuArgTyrAlaThrIleLeu	146
DB	355	GCTATGCTTTTGACCGTTATGTGGCACTCTGTGCTCCACTACATTACGCAACCATCTTG	414

genomic survey sequence.

AY401478

AY401478.1 GI:39757467

GSS.

Homo sapiens (human)

ORGANISM

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE

AUTHORS

1 (bases 1 to 936)
Clark, A.G., Gnanowski, S., Nielson, R., Thomas, P., Kejariwal, A., Todd, M.A., Tanenbaum, D.M., Civello, D.R., Lu, F., Murphy, B., Ferreria, S., Wang, G., Zheng, X.H., White, T.J., Shinsky, J.J., Adams, M.D. and Cargill, M.

TITLE

Inferring nonneutral evolution from human-chimp-mouse orthologous gene trios

JOURNAL

PUBMED

Science 302 (5652), 1960-1963 (2003)

REFERENCE

AUTHORS

2 (bases 1 to 936)
Clark, A.G., Gnanowski, S., Nielson, R., Thomas, P., Kejariwal, A., Todd, M.A., Tanenbaum, D.M., Civello, D.R., Lu, F., Murphy, B., Ferreria, S., Wang, G., Zheng, X.H., White, T.J., Shinsky, J.J., Adams, M.D. and Cargill, M.

TITLE

Direct Submission

JOURNAL

COMMENT

Submitted (16-NOV-2003) Celera Genomics, 45 West Gude Drive, Rockville, MD 20850, USA

FEATURES

source

1..936

/organism="Homo sapiens"

/mol_type="genomic DNA"

/db_xref="taxon:9606"

<1..>936

/locus_tag="HCM0904"

ORIGIN

Alignment Scores:

Pred. No.: 4,05e-93

Score: 904.00

Percent Similarity: 71.70%

Best Local Similarity: 54.34%

Query Match: 52.62%

DB: 29

Length: 936

Matches: 169

Conservative: 54

Mismatches: 82

Indels: 6

Gaps: 2

US-10-081-775-2 (1-329) x AY401478 (1-936)

QY 7 HisAsnMetGluSerProHisHisThrAspValAspProSerValPhePheLeuLeuGly 26

DB 10 CACAAACAGACATATTTTAC-----CCAGTCACATTTTCTCTCATGGA 54

QY 27 IleProGlyLeuGluGlnPheHisLeuTrpLeuSerLeuProValCysGlyLeuGlyThr 46

DB 55 ATCCAGGCTCTGGAAGACTTCCACATGTGGATCTCCGGGCTTCTGCTCTGTACCTT 114

QY 47 AlaThrIleValGlyAsnIleThrIleLeuValValAlaThrGluProValLeuHis 66

DB 115 GTGCTTGTGGCAATGTCACCATCTCTGTAGTATCATCAAGGTAGAACAGACTCTCGG 174

QY 67 LysProValTyrLeuPheLeuCysMetLeuSerThrIleAspLeuAlaAlaSerValSer 86

DB 175 GAGCCCAATGTTCTCTCTGGCCATCTTTTCCACTATTCATTTGGGCTTCTCTGCAACC 234

QY 87 ThrValProLysLeuLeuAlaIlePheTrpCysGlyAlaGlyHisIleSerAlaSerAla 106

DB 235 TC GTGCTCGCATGTGGGTATCTCTGTTGATGCTCAGAGATTAATATGAGGCT 294

QY 107 CysLeuAlaHisMetPhePheIleHisAlaPheCysMetMetGluSerThrValLeuLeu 126

DB 295 TGTGTGCCCGAGATGTTTCTGATCCATGCTTCACTGGCATGGGCTGAGGCTTACTG 354

QY 127 AlaMetAlaPheAspArgTyrValAlaIleCysHisProLeuArgTyrAlaThrIleLeu 146

DB 355 GCTATGCTTTTGACCGTTATGTGGCACTCTGTGCTCCACTACATTACGCAACCATCTTG 414

QY	Db	309	IleArgLysArgValValArgValPheGlnSerGly	320	330
QY	Db	898	ATCTGGGATGGGGCCCTCCGGCTTCTGAAGTGGGGC	933	943
RESULT 5					
AY414225					
LOCUS					
DEFINITION					
ACCESSION					
VERSION					
KEYWORDS					
SOURCE					
ORGANISM					
REFERENCE					
AUTHORS					
TITLE					
JOURNAL					
PUBMED					
REFERENCE					
AUTHORS					
TITLE					
JOURNAL					
PUBMED					
REFERENCE					
AUTHORS					
TITLE					
JOURNAL					
PUBMED					
REFERENCE					
AUTHORS					
TITLE					
JOURNAL					
PUBMED					
REFERENCE					
AUTHORS					
TITLE					
JOURNAL					
PUBMED					
REFERENCE					
AUTHORS					
TITLE					
JOURNAL					
PUBMED					
REFERENCE					
AUTHORS					
TITLE					
JOURNAL					
PUBMED					
REFERENCE					
AUTHORS					
TITLE					
JOURNAL					
PUBMED					
REFERENCE					
AUTHORS					
TITLE					
JOURNAL					
PUBMED					
REFERENCE					
AUTHORS					
TITLE					
JOURNAL					
PUBMED					
REFERENCE					
AUTHORS					
TITLE					
JOURNAL					
PUBMED					
REFERENCE					
AUTHORS					
TITLE					
JOURNAL					
PUBMED					
REFERENCE					
AUTHORS					
TITLE					
JOURNAL					
PUBMED					
REFERENCE					
AUTHORS					
TITLE					
JOURNAL					
PUBMED					
REFERENCE					
AUTHORS					
TITLE					
JOURNAL					
PUBMED					
REFERENCE					
AUTHORS					
TITLE					
JOURNAL					
PUBMED					
REFERENCE					
AUTHORS					
TITLE					
JOURNAL					
PUBMED					
REFERENCE					
AUTHORS					
TITLE					
JOURNAL					
PUBMED					
REFERENCE					
AUTHORS					
TITLE					
JOURNAL					
PUBMED					
REFERENCE					
AUTHORS					
TITLE					
JOURNAL					
PUBMED					
REFERENCE					
AUTHORS					
TITLE					
JOURNAL					
PUBMED					
REFERENCE					
AUTHORS					
TITLE					
JOURNAL					
PUBMED					
REFERENCE					
AUTHORS					
TITLE		</			

Qy	309	IleArgGlyeArgValValArgValPheGlnSerGly	320
Db	898	ATCTGGGATGGGGCCCTCGGCTTCTGAAGTGCGGC	933
RESULT 5			
AY414225			
LOCUS			
DEFINITION	Homo sapiens HCM5153 gene, VIRTUAL TRANSCRIPT, partial sequence, genomic survey sequence.		
ACCESSION	AY414225		
VERSION	AY414225.1	GI:39770187	
KEYWORDS	GSS.		
SOURCE	Homo sapiens (human)		
ORGANISM	Homo sapiens		
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo. 1 (bases 1 to 957) Clark,A.G., Gnanowski,S., Nielson,R., Thomas,P., Kejariwal,A., Todd,M.A., Tanenbaum,D.M., Civello,D.R., Lu,F., Murphy,B., Perriera,S., Wang,G., Zheng,X.H., White,T.J., Sninsky,J.J., Adams,M.D. and Cargill,M. Inferring nonneutral evolution from human-chimp-mouse orthologous gene trios Science 302 (5652), 1960-1963 (2003) 14671302		
AUTHORS	Clark,A.G., Gnanowski,S., Nielson,R., Thomas,P., Kejariwal,A., Todd,M.A., Tanenbaum,D.M., Civello,D.R., Lu,F., Murphy,B., Perriera,S., Wang,G., Zheng,X.H., White,T.J., Sninsky,J.J., Adams,M.D. and Cargill,M.		
TITLE	Inferring nonneutral evolution from human-chimp-mouse orthologous gene trios		
JOURNAL	Science 302 (5652), 1960-1963 (2003)		
PUBMED	14671302		
REFERENCE	2 (bases 1 to 957) Clark,A.G., Gnanowski,S., Nielson,R., Thomas,P., Kejariwal,A., Todd,M.A., Tanenbaum,D.M., Civello,D.R., Lu,F., Murphy,B., Perriera,S., Wang,G., Zheng,X.H., White,T.J., Sninsky,J.J., Adams,M.D. and Cargill,M. Direct Submission Submitted (16-NOV-2003) Celera Genomics, 45 West Gude Drive, Rockville, MD 20850, USA This sequence was made by sequencing genomic exons and ordering them based on alignment. Location/Qualifiers 1..957 /organism="Homo sapiens" /mol_type="genomic DNA" /db_xref="taxon:9606" <1..>957 /locus_tag="HCM5153"		
AUTHORS	Clark,A.G., Gnanowski,S., Nielson,R., Thomas,P., Kejariwal,A., Todd,M.A., Tanenbaum,D.M., Civello,D.R., Lu,F., Murphy,B., Perriera,S., Wang,G., Zheng,X.H., White,T.J., Sninsky,J.J., Adams,M.D. and Cargill,M. Direct Submission Submitted (16-NOV-2003) Celera Genomics, 45 West Gude Drive, Rockville, MD 20850, USA This sequence was made by sequencing genomic exons and ordering them based on alignment. Location/Qualifiers 1..957 /organism="Homo sapiens" /mol_type="genomic DNA" /db_xref="taxon:9606" <1..>957 /locus_tag="HCM5153"		
COMMENT	This sequence was made by sequencing genomic exons and ordering them based on alignment. Location/Qualifiers 1..957 /organism="Homo sapiens" /mol_type="genomic DNA" /db_xref="taxon:9606" <1..>957 /locus_tag="HCM5153"		
FEATURES	source		
gene			
ORIGIN			
Alignment Scores:			
Pred. No.:	1.97e-91	Length:	957
Score:	889.50	Matches:	160
Percent Similarity:	69.41%	Conservative:	51
Best Local Similarity:	52.63%	Mismatches:	92
Query Match:	51.78%	Indels:	1
DB:	29	Gaps:	1
US-10-081-775-2 (1-329) x AY414225 (1-957)			
Qy	19	ProSerValPhePheLeuLeuGlyIleProGlyLeuGluGlnPheHisLeuTrpLeuSer	38
Db	34	CCAGACACTTCTTCTTTAAACGGGATCCAGGGGTGGAGGCTGCCCACTTCGATTGGC	93
Qy	39	LeuProValCysGlyLeuGlyThrAlaThrIleValGlyAsnIleThrLeuValVal	58
Db	94	ATCCCTTTCTGTGCATGATCTTTGTAGCACTGGTTGGAATGTCGCTCATCTCGTC	153
Qy	59	VallAlaThrGluProValLeuHisLysProValTyrrLeuPheLeuCysMetLeuSerThr	78
Db	154	ATTGCCATGGCAATGCTCTTCATGCACCTATGACCTTCCTCTGCTCTCTCTCACATC	213
Qy	79	IleAspLeuAlaAalaSerValSerThrValProlysLeuLeuAlaIlePheTrpCysGly	98
Db	214	ACAGACCTGGCTCTCAGTCTTACCACCTGTGCCAAGATGCTGCCATTTTGGCTCCAT	273
Qy	99	AlaGlyHisIleSerAlaSerAlaCysLeuAlaHisMetPhePheIleHisAlaPheCys	118

JOURNAL Submitted (16-NOV-2003) Celera Genomics, 45 West Gude Drive, Rockville, MD 20850, USA
COMMENT This sequence was made by sequencing genomic exons and ordering them based on alignment.
FEATURES Location/Qualifiers
source 1..954
/organism="Mus musculus"
/mol_type="genomic DNA"
/db_xref="taxon:10090"
<1..>954
gene /locus_tag="Hmw5153"
ORIGIN
Alignment Scores:
Pred. No.: 1.26e-90 Length: 954
Score: 882.50 Matches: 162
Percent Similarity: 70.59% Conservative: 54
Best Local Similarity: 52.94% Mismatches: 89
Query Match: 51.37% Indels: 1
DB: 29 Gaps: 1
US-10-081-775-2 (1-329) x AY414227 (1-954)
Qy 19 ProSerValPhePheLeuGlyLeuProGlyLeuGlnPheHisLeuTyrLeuSer 38
Db 31 CCAGTCATCTCTTCTTAAGTCCAGGCTGGAGTTCGCCACCTCTGGATTGCC 90
Qy 39 LeuProValCysGlyLeuGlyThrAlaThrIleValGlyAsnIleThrIleLeuVal 58
Db 91 ATCCCTTCTGTGTCATGTATGCTGTCTGCTGCTGGATGCTGCCCTCATCTTATC 150
Qy 59 ValAlaThrGluProValLeuHisProValTyrLeuPheLeuCysMetLeuSerThr 78
Db 151 ATTGGCAGAGAGGTGCTCCACACCCCATGTACCTTCTCTGCTGCTCTCTCATCTC 210
Qy 79 IleAspLeuAlaAlaSerValSerThrValProLysLeuLeuAlaIlePheTyrCysGly 98
Db 211 ACTGACTTGGCTCTCAGCTCCACCTGTACCTAAAGTTGGCCATCTATGGCTCCAT 270
Qy 99 AlaGlyHisIleSerAlaSerAlaCysLeuAlaHisMetPheIleHisAlaPheCys 118
Db 271 TCTAATGAGATTCTCTTGGTGGATGCTCCGCCAGATGTTTGTGTCCACTCCATCTAT 330
Qy 119 MetMetGluSerThrValLeuLeuAlaMetAlaPheAspArgTyrValAlaIleCysHis 138
Db 331 GCTCTGTAGTCTCGGTCTCTTCCATGGCTTGTATCGATATGTAGTATCTGCAAC 390
Qy 139 ProLeuArgTyrAlaThrIleLeuThrAspThrIleAlaHisIleGlyValAlaAla 158
Db 391 CCACCTGAGTATACACCATCTCTCAACCATATGTATAGCCCAATATCTTTGCTGGG 450
Qy 159 ValValArgGlySerLeuLeuMetLeuProCysProPhePheIleGlyArgLeuAsnPhe 178
Db 451 ATAGTCCGTAGTGTGGCTATGTCTCCCTTTTCATCTTTTGTGAGGCGACTGCCCTAC 510
Qy 179 CysGlnSerHisValIleLeuHisThrTyrCysGluHisMetAlaValValLysLeuAla 198
Db 511 TGTGGTCCCGTGCATGACTCACATATGTGAGCATATGGCATTCCTAGGCTTGGC 570
Qy 199 CysGlyAspThrArgProAsnArgValTyrGlyLeuThrAlaAlaLeuValIleGly 218
Db 571 TGTGCCAATATTAATGTTATATTGTTATGGTTGACTGTAGCTTTGTGGCCATGGC 630
Qy 219 ValAspLeuPheCysIleGlyLeuSerTyrAlaLeuIleAlaGlnAlaValLeuArgLeu 238
Db 631 CTGGATTCCATCTCATGCTGCTATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 690
Qy 239 SerSerHisGluAlaArgSerIysAlaLeuGlyThrCysGlySerHisValCysValIle 258
Db 691 CCATCCCGTATGCTCAACAAAGGCTGTGATGCTGATGCTGATGCTGATGCTGAT 750
Qy 259 LeuIleSerTyrThrProAlaLeuPheSerPheThrHisArgPheGlyHisHis 277

Db 751 CTGCTTTTATATACCTGCTTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCA 810
Qy 278 ValProValHisIleHisIleLeuLeuAlaAsnValTyrLeuLeuLeuProAlaLeu 297
Db 811 GTTCCCAAGCATGTGCATATCTCTCTGCTAACCTCTATGTGCTGGTGGCTCCCGTGC 870
Qy 298 AsnProValValTyrGlyValLysGlnIleArgLysArgValValArgValPhe 317
Db 871 AATCCCATCATCTACGTGCGAAGAACCAAGAGATCGGATCGGCTCTTAAACTGCTT 930
Qy 318 GlnSerGlyGlnGlyMet 323
Db 931 CACTTGGGAAGGACTTA 948
RESULT 7
AY401480 936 bp DNA linear GSS 12-DEC-2003
LOCUS Mus musculus HCM0904 gene, VIRTUAL TRANSCRIPT, partial sequence,
DEFINITION genomic survey sequence.
ACCESSION AY401480
VERSION AY401480.1 GI:39757469
KEYWORDS GSS.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1 (bases 1 to 936)
AUTHORS Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejariwal,A.,
Todd,M.A., Tanenbaum,D.M., Civello,D.R., Lu,F., Murphy,B.,
Ferriera,S., Wang,G., Zheng,X.H., White,T.J., Sninsky,J.J.,
Adams,M.D. and Cargill,M.
TITLE Inferring nonneutral evolution from human-chimp-mouse orthologous
gene trios
JOURNAL Science 302 (5652), 1960-1963 (2003)
PUBMED 14671302
REFERENCE 2 (bases 1 to 936)
AUTHORS Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejariwal,A.,
Todd,M.A., Tanenbaum,D.M., Civello,D.R., Lu,F., Murphy,B.,
Ferriera,S., Wang,G., Zheng,X.H., White,T.J., Sninsky,J.J.,
Adams,M.D. and Cargill,M.
TITLE Direct Submission
JOURNAL Submitted (16-NOV-2003) Celera Genomics, 45 West Gude Drive,
Rockville, MD 20850, USA
COMMENT This sequence was made by sequencing genomic exons and ordering
them based on alignment.
FEATURES Location/Qualifiers
source 1..936
/organism="Mus musculus"
/mol_type="genomic DNA"
/db_xref="taxon:10090"
<1..>936
/locus_tag="HCM0904"
gene
ORIGIN
Alignment Scores:
Pred. No.: 2.08e-90 Length: 936
Score: 880.50 Matches: 163
Percent Similarity: 70.42% Conservative: 56
Best Local Similarity: 52.41% Mismatches: 87
Query Match: 51.25% Indels: 5
DB: 29 Gaps: 2
US-10-081-775-2 (1-329) x AY401480 (1-936)
Qy 7 HisAsnMetGluSerProHisHisThrAspValAspProSerValPhePheLeuGly 26
Db 7 CATAATAATAGAGCGTTTCTCAC-----CCAGATACATTTTCTCATGGG 54
Qy 27 IleProGlyLeuGluGlnPheHisLeuTyrLeuSerLeuProValCysGlyLeuGlyThr 46
Db 55 ATCCCAAGTCTGAGGAATATCCATGCTGATTTCTGCTTCTGCTGCTGCTGCTGCTT 114
Qy 47 AlaThrIleValGlyAsnIleThrIleLeuValValAlaThrGluProValLeuHis 66

CDS
 /dev stage="16 days neonate"
 348..1310
 /note="unnamed protein product; OLFACTORY RECEPTOR MOR18-2
 (SPTR|AAL60674, evidence: FASTV, 100%ID, 100%length,
 match=960)
 putative"
 /codon_start=1
 /protein_id="BAC29396.1"
 /db_xref="GI:26331332"
 /translation="MSSCNFTHTAFLIGFLEAEHFWGFPILLSMYAVAFNGCIV
 VFVITERSLHAPMYLFCMLAIDLALSTMPKILALFPWDSREITFDCLAQMF
 IHTLSAEHSTILLAMAPRYVAICPLRHAALVNTVQIGMVALVRSGLFPPLPL
 LIKALFCHSVNLSHCYHODVMKLAITDPLNVVYGLTALLVMGVDMFISLSYF
 LIIRVLQPSKERAKAFVCVSHISVLAIFYPLIGLSVYVHRFGNSLDPIRVHVMG
 DVYLLPVPVPIIYGAKTKQIRLVLMFKISKDKLEAGNT"
 1942..1947
 /note="putative"
 1964
 /note="putative"

polyA_signal

polyA_site

ORIGIN

Alignment Scores:

Pred. No.: 1.8e-88 Length: 1964
 Score: 868.00 Matches: 166
 Percent Similarity: 67.18% Conservative: 53
 Best Local Similarity: 50.92% Mismatches: 107
 Query Match: 50.52% Indels: 0
 DB: 11 Gaps: 0

US-10-081-775-2 (1-329) x AK036356 (1-1964)

QY 2 SerSerThrLeuGlyHisAsnMetGluSerProHisHisThrAspValAspProSerVal 21
 DB 318 TCTTCCAGTCCGCCCACTTGTGTCACGATGAGTTCGCACTTCCATCCATCCACC 377
 QY 22 PhePheLeuGlyIleProGlyLeuGluInPheHisLeuThrLeuSerLeuProVal 41
 DB 378 TTCTGCTTATGGTATCCAGACTGGAGAGCTCACTTTGGTGGCTTCCCTCG 437
 QY 42 CysGlyLeuGlyThrAlaThrIleValGlyAsnIleThrIleLeuValValAlaThr 61
 DB 438 CTTTCCATGTATGCTGTAGCACTGTTTGGAACTGCATTTGCTTTCATAGTGAACA 497
 QY 62 GluProValLeuHisGlyProValTyLeuPheLeuCysMetLeuSerThrIleAspLeu 81
 DB 498 GAGCGAGCCTGATGACCCCACTGACCTTTTCTGTCATGCTGCACCTATTGATCG 557
 QY 82 AlaAlaSerValSerThrValProGlyLeuAlaIlePheThrCysGlyAlaGlyHis 101
 DB 558 GTTTTGTCCACATCCACATGCCAAGATTCCTGCGCTTTTGGTTGACTCCCGGAG 617
 QY 102 IleSerAlaSerAlaCysLeuAlaHisMetPhePheIleHisAlaPheCysMetMetGlu 121
 DB 618 ATTACTTTTGTGCTGCTGCTGCCAGATGTTTTCATTCATCACTCTCTCAGCAATGAA 677
 QY 122 SerThrValLeuLeuAlaMetAlaPheAspArgTyValAlaIleCysHisProLeuArg 141
 DB 678 TCTACTATCTGCTGCCATGGCTTTGACCGTACGTGGCTATCTGCCACCCACTGCT 737
 QY 142 TyrAlaThrIleLeuThrAspThrIleAlaHisIleGlyValAlaAlaValValArg 161
 DB 738 CAGCGTGTCTCTCAACAATACAGTACAGTCCAAATAGGCATGGTGGCTCTGGTCCG 797
 QY 162 GlySerLeuLeuMetLeuProCysProPheIleGlyArgLeuAsnPheCysGlnSer 181
 DB 798 GGATCCCTATTCTTTTCCACTCCCACTACTGATCAAGCGACTGGCTTCTGTACTCC 857
 QY 182 HisValIleLeuHisThrTyCysGluHisMetAlaValValIleLeuAlaCysGlyAsp 201
 DB 858 AATGTGCTCTCCACTCTTATGTGTGCCACGAGTGTGATGAGTTGCGCTTATACAGAC 917
 QY 202 ThrArgProAsnArgValTyGlyLeuThrAlaAlaLeuLeuValIleGlyValAspLeu 221

Db 918 ACATTACCCAATAGTAGTCTATGGTCTAACTGATGCCATTCTGTAGTATGGGTGTAGATGTC 977
 QY 222 PheCysIleGlyLeuSerTyAlaLeuIleAlaGlnAlaValLeuArgLeuSerSerHis 241
 Db 978 ATGTTATCTCTCTGCTCTACTTCTGATCATATAGACGGTCTGCAACTGCTTCCAAG 1037
 QY 242 GluAlaArgSerLysAlaLeuGlyThrCysGlySerHisValCysValIleLeuIleSer 261
 Db 1038 TCTGAGCGAGTAAAGCATTTGGGACTTGTGTGCACATAGTAGTGTGCTGCTGCTTC 1097
 QY 262 TyrThrProAlaLeuPheSerPhePheThrHisArgPheGlyHisHisValProValHis 281
 Db 1098 TATGTACCACTCATTTGGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1157
 QY 282 IleHisIleLeuLeuAlaAsnValTyLeuLeuLeuProProAlaLeuAsnProValVal 301
 Db 1158 GTGCATGTTCTAATGGGGGATGTCTACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1217
 QY 302 TyrGlyValLysThrLysGlnIleArgLysArgValValArgValPheGlnSerGlyGln 321
 Db 1218 TACGGTGTAAAGCAACACATCAGAACACGGGTCTTGGCTATGTTTCAAGATCAGCTGT 1277
 QY 322 GlyMetGlyIleLysAla 327
 Db 1278 GACAAGGACATTGAAGCT 1295
 RESULT 9
 AK028467
 LOCUS
 DEFINITION
 AK028467
 3410 bp mRNA linear HTC 18-SEP-2003
 Mus musculus 0 day neonate skin cDNA, RIKEN full-length enriched
 library, clone:4631413D08 product:OLFACTORY RECEPTOR MOR18-2, full
 insert sequence.
 ACCESSION
 VERSION
 KEYWORDS
 SOURCE
 ORGANISM
 AK028467.1 GI:26324423
 HTC; CAP trapper.
 Mus musculus (house mouse)
 Mus musculus
 Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 1
 Carninci, P. and Hayashizaki, Y.
 High-efficiency full-length cDNA cloning
 Meth. Enzymol. 303, 19-44 (1999)
 99279253
 MEDLINE
 PUBMED
 10349636
 2
 Carninci, P., Shibata, Y., Hayateu, N., Sugahara, Y., Shibata, K.,
 Itoh, M., Konno, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.
 Normalization and subtraction of cap-trapper-selected cDNAs to
 prepare full-length cDNA libraries for rapid discovery of new genes
 Genome Res. 10 (10), 1617-1630 (2000)
 20499374
 MEDLINE
 PUBMED
 11042159
 3
 Shibata, K., Itoh, M., Aizawa, K., Nagaoka, S., Sasaki, N., Carninci, P.,
 Konno, H., Akiyama, J., Nishi, K., Kitsuai, T., Tashiro, H., Itoh, M.,
 Sumi, N., Ishii, Y., Nakamura, S., Hazama, M., Nishine, T., Harada, A.,
 Yamamoto, R., Matsumoto, H., Sakaguchi, S., Ikegami, T., Kashiwagi, K.,
 Fujiwara, S., Inoue, K., Togawa, Y., Izawa, M., Ohara, E., Watahiki, M.,
 Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, I., Matsura, S., Kawai, J.,
 Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A. and Hayashizaki, Y.
 RIKEN integrated sequence analysis (RISA) system--384-format
 sequencing pipeline with 384 multicapillary sequencer
 Genome Res. 10 (11), 1757-1771 (2000)
 20530913
 MEDLINE
 PUBMED
 11076861
 4
 The RIKEN Genome Exploration Research Group Phase II Team and the
 FANTOM Consortium.
 Functional annotation of a full-length mouse cDNA collection
 Nature 409, 685-690 (2001)
 5
 The FANTOM Consortium and the RIKEN Genome Exploration Research


```

DEFINITION Mus musculus OR52R1 gene, VIRTUAL TRANSCRIPT, partial sequence,
genomic survey sequence.
ACCESSION AY410598
VERSION AY410598.1 GI:39766568
KEYWORDS GSS.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
REFERENCE
AUTHORS Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejariwal,A.,
Todd,M.A., Tanenbaum,D.M., Civello,D.R., Lu,F., Murphy,B.,
Ferriera,S., Wang,G., Zheng,X.H., White,T.J., Shinsky,J.J.,
Adams,M.D. and Cargill,M.
TITLE Inferring nonneutral evolution from human-chimp-mouse orthologous
gene trios
JOURNAL Science 302 (5652), 1960-1963 (2003)
PUBMED 14671302
REFERENCE
AUTHORS Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejariwal,A.,
Todd,M.A., Tanenbaum,D.M., Civello,D.R., Lu,F., Murphy,B.,
Ferriera,S., Wang,G., Zheng,X.H., White,T.J., Shinsky,J.J.,
Adams,M.D. and Cargill,M.
TITLE Direct Submission
JOURNAL Submitted (16-NOV-2003) Celera Genomics, 45 West Gude Drive,
Rockville, MD 20850, USA
COMMENT This sequence was made by sequencing genomic exons and ordering
them based on alignment.
FEATURES
source
1..948
/organism="Mus musculus"
/mol_type="genomic DNA"
/db_xref="taxon:10090"
<1..>948
/gene="OR52R1"
/locus_tag="HMC3944"
gene
ORIGIN
Alignment Scores:
Pred. No.: 6,698-89 Length: 948
Score: 867.50 Matches: 168
Percent Similarity: 67.0% Conservative: 42
Best Local Similarity: 53.6% Mismatches: 98
Query Match: 50.4% Indels: 5
DB: 29 Gaps: 1
US-10-081-775-2 (1-329) x AY410600 (1-948)
QY 5 LeuGlyHisAsnMetGluSerProHisThrAspValAspProSerValPheLeu 24
Db 7 TTGGCTTCAGGAACAGCTCTCCCACTCTACA-----TTTTTCATCTTG 51
QY 25 LeuGlyIleProGlyLeuGluGlnPheHisLeuTrpLeuSerLeuProValCysGlyLeu 44
Db 52 CTTGAATCCAGGACTGGAGAATTACCAATTCGGTTGCTTCCATCTGTGTGTCATG 111
QY 45 GlyThrAlaThrIleValGlyAsnIleThrIleLeuValValAlaThrGluProVal 64
Db 112 TATATTGGCGCAGTCAGTGGAAACATCACCATCTTCATATATCCGAATGACCACACC 171
QY 65 LeuHisLysProValTrpLeuPheLeuCysMetLeuSerThrIleAspLeuAlaAlaSer 84
Db 172 CTACATGAGCCCATGTATCTCTTCTGGCGATGCTGGCTACCACTGACCTTGTCTCTG 231
QY 85 ValSerThrValProLysLeuLeuAlaIlePheThrCysGlyAlaGlyHisIleSerAla 104
Db 232 TCTCTCCACACACCCAAAATGCTGGCTATACCTGCTTTTCATGATCATGAAGATTGAAT 291
QY 105 SerAlaCysLeuAlaHisMetPhePheIleHisAlaPheCysMetGluSerThrVal 124
Db 292 CATGCTGCTCTATCCAGGTGTCTTCATACATGCTCTTCTCTGCTGGAGTCTGGGGTG 351
QY 125 LeuLeuAlaMetAlaPheAspArgTrpValAlaIleCysHisProLeuArgTrpAlaThr 144

```

```

Db 352 CTCATGACCATGGCTTGGACCGTTATGTGCTATCTCTTCCCACTCCGACATTCACGC 411
QY 145 IleLeuThrAspThrIleIleAlaHisIleGlyValAlaAlaValValArgGlySerLeu 164
Db 412 ATCTCCACCATCTGCAGTCATCAAACTAGTGCTGTGTGTGTGTGTGTGTGTGTGTGTG 471
QY 165 LeuMetLeuProCysProPhePheIleGlyArgLeuAsnPheCysGlnSerHisValIle 184
Db 472 TGGGTGAGCCCTTCTGTTCATGCTTCCAGGATGCCCTTCTGCCCCCAACAAGTCAAT 531
QY 185 LeuHisThrTrpCysGluHisMetAlaValValLysLeuAlaCysGlyAspThrArgPro 204
Db 532 CCCAGTCTCTACTGTGACACATGGCTGTCTTAAGTTGGTATGTGTGTGTGTGTGTGTG 591
QY 205 AsnArgValTrpGlyLeuThrAlaLeuLeuValIleGlyValAlaAspLeuPheCysIle 224
Db 592 AATCGTGATATGGACTCTTTGTGCTTCTCTCTGTGTGTGTGTGTGTGTGTGTGTGTG 651
QY 225 GlyLeuSerTrpAlaLeuIleAlaGlnAlaValLeuArgLeuSerSerHisGluAlaArg 244
Db 652 AGTGATCTTATGATGATGATCTGTGAGACTGTGTGTGTGTGTGTGTGTGTGTGTGTG 711
QY 245 SerLysAlaLeuGlyThrCysGlySerHisValCysValIleLeuIleSerTrpThrPro 264
Db 712 CTCAGGCTTTTGGTACATGTCTCTCATATTGGTGTGCATATTAACTTATATATATCCCA 771
QY 265 AlaLeuPheSerPhePheThrHisArgPheGlyHisHisValProValHisIleHisIle 284
Db 772 GCCCTTTTCACCTTCTCCTCACCACCGCTTGTGTGTGTGTGTGTGTGTGTGTGTGTG 831
QY 285 LeuLeuAlaAsnValTrpLeuLeuLeuProAlaLeuAsnProValValTrpGlyVal 304
Db 832 ATGTTTGGCAATGCTATCTTCTAGTCTCTCTATGCTCAATCCCATCATATATGAGTT 891
QY 305 LysThrLysGlnIleArgLysArgValValArgValPhe 317
Db 892 AGAACCAACAGATCAGACAGAGGTTTACCCGAGGGTTT 930

```

```

RESULT 11
AY410598 948 bp DNA linear GSS 16-DEC-2003
LOCUS Homo sapiens OR52R1 gene, VIRTUAL TRANSCRIPT, partial sequence,
DEFINITION genomic survey sequence.
ACCESSION AY410598
VERSION AY410598.1 GI:39766566
KEYWORDS GSS.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
REFERENCE
AUTHORS Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejariwal,A.,
Todd,M.A., Tanenbaum,D.M., Civello,D.R., Lu,F., Murphy,B.,
Ferriera,S., Wang,G., Zheng,X.H., White,T.J., Shinsky,J.J.,
Adams,M.D. and Cargill,M.
TITLE Inferring nonneutral evolution from human-chimp-mouse orthologous
gene trios
JOURNAL Science 302 (5652), 1960-1963 (2003)
PUBMED 14671302
REFERENCE
AUTHORS Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejariwal,A.,
Todd,M.A., Tanenbaum,D.M., Civello,D.R., Lu,F., Murphy,B.,
Ferriera,S., Wang,G., Zheng,X.H., White,T.J., Shinsky,J.J.,
Adams,M.D. and Cargill,M.
TITLE Direct Submission
JOURNAL Submitted (16-NOV-2003) Celera Genomics, 45 West Gude Drive,
Rockville, MD 20850, USA
COMMENT This sequence was made by sequencing genomic exons and ordering
them based on alignment.
FEATURES
source
1..948
/organism="Homo sapiens"

```


source

```
1. 1086
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/map="11p15.5"
/clone="PTWB298.21SP6"
/clone_lib="Human"
```

ORIGIN

Alignment Scores:

```
Pred. No.:      2.95e-81      Length:      1086
Score:          802.00      Matches:      151
Percent Similarity: 66.83%      Conservative: 49
Best Local Similarity: 50.50%      Mismatches: 99
Query Match:      46.68%      Indels:      0
DB:              28          Gaps:        0
```

US-10-081-775-2 (1-329) x AF101706 (1-1086)

```
QY 21 ValPhePheLeuLeuGlyLeuProGlyLeuGluGlnPheHisLeuTyrLeuSerLeuPro 40
Db 21 ::::: ||| ::::: ||| ::::: ||| ::::: ||| ::::: ||| ::::: ||| ::::: |||
QY 41 ValCysGlyLeuGlyThrAlaThrIleValGlyAsnIleThrIleLeuValValAla 60
Db 41 ::::: ||| ::::: ||| ::::: ||| ::::: ||| ::::: ||| ::::: |||
QY 61 ThrGluProValLeuHisLysProValTyrLeuPheLeuCysMetLeuSerThrIleAsp 80
Db 61 ::::: ||| ::::: ||| ::::: ||| ::::: ||| ::::: ||| ::::: |||
QY 81 LeuAlaAlaSerValSerThrValProLysLeuLeuAlaIlePheTyrCysGlyAlaGly 100
Db 81 ::::: ||| ::::: ||| ::::: ||| ::::: ||| ::::: ||| ::::: |||
QY 101 HisIleSerAlaSerAlaCysLeuAlaHisMetPhePheIleHisAlaPheCysMetMet 120
Db 101 ::::: ||| ::::: ||| ::::: ||| ::::: ||| ::::: ||| ::::: |||
QY 121 GluSerThrValLeuLeuAlaMetAlaPheAspArgTyrValAlaIleCysHisProLeu 140
Db 121 ::::: ||| ::::: ||| ::::: ||| ::::: ||| ::::: ||| ::::: |||
QY 141 ArgTyrAlaThrIleLeuThrAspThrIleAlaHisIleGlyValAlaAlaValVal 160
Db 141 ::::: ||| ::::: ||| ::::: ||| ::::: ||| ::::: ||| ::::: |||
QY 161 ArgGlySerLeuLeuMetLeuProCysProPhePheIleGlyArgLeuAsnPheCysGln 180
Db 161 ::::: ||| ::::: ||| ::::: ||| ::::: ||| ::::: ||| ::::: |||
QY 181 SerHisValIleLeuHisThrTyrCysGluHisMetAlaValValLysLeuAlaCysGly 200
Db 181 ::::: ||| ::::: ||| ::::: ||| ::::: ||| ::::: ||| ::::: |||
QY 201 AspThrArgProAsnArgValTyrGlyLeuThrAlaAlaLeuLeuValIleGlyValAsp 220
Db 201 ::::: ||| ::::: ||| ::::: ||| ::::: ||| ::::: ||| ::::: |||
QY 221 LeuPheCysIleGlyLeuSerTyrAlaLeuIleAlaGlnAlaValLeuArgLeuSerSer 240
Db 221 ::::: ||| ::::: ||| ::::: ||| ::::: ||| ::::: ||| ::::: |||
QY 241 HisGluAlaArgSerLysAlaLeuGlyThrCysGlySerHisValCysValIleLeuIle 260
Db 241 ::::: ||| ::::: ||| ::::: ||| ::::: ||| ::::: ||| ::::: |||
QY 261 CCAGGAAGAGAGAGAGAACCTTCAGTACATGTGTCTCCACACATGTGGCTTTTGTCTATA 280
Db 261 ::::: ||| ::::: ||| ::::: ||| ::::: ||| ::::: ||| ::::: |||
QY 281 SerTyrThrProAlaLeuPheSerPhePheThrHisArgPheGlyHisHisValProVal 300
Db 281 ::::: ||| ::::: ||| ::::: ||| ::::: ||| ::::: ||| ::::: |||
QY 301 ValTyrGlyValLysThrLysGlnIleArgLysArgValValArgValPheGlnSer 319
Db 301 ::::: ||| ::::: ||| ::::: ||| ::::: ||| ::::: ||| ::::: |||
QY 81 ATCTACAGTGTGAAGAACCAACAGATACGTAGAGCTGTGATAAAATTTCTCCATTCC 25
Db 81 ::::: ||| ::::: ||| ::::: ||| ::::: ||| ::::: ||| ::::: |||
```

Search completed: August 27, 2004, 20:56:13
Job time : 3054 secs

```
Db 141 TATGTACATACCTATGATTGCTAAACACCTACCTGCTGACCTCCCTTTTGATGAACCTGTC 82
QY 301 ValTyrGlyValLysThrLysGlnIleArgLysArgValValArgValPheGlnSer 319
Db 81 ATCTACAGTGTGAAGAACCAACAGATACGTAGAGCTGTGATAAAATTTCTCCATTCC 25
```